



Facultad de Ciencias Económicas y Empresariales

Predicción de la Generación de Electricidad a partir del Biogás en Galicia: Análisis de Tendencias y Proyección a Corto Plazo

Autor: Candela Sánchez-Toscano Botas

Director: Jenny Alexandra Cifuentes Quintero

MADRID | Junio 2026

Resumen

La predicción de la generación eléctrica a partir de biogás constituye una herramienta relevante para la planificación operativa de plantas de digestión anaerobia y para la integración de esta tecnología en sistemas eléctricos con una elevada participación de fuentes renovables. A diferencia de recursos como la energía eólica o solar, la variabilidad del biogás no depende principalmente de las condiciones meteorológicas, sino de factores operativos y biológicos vinculados al tipo de sustrato, las condiciones de digestión, la gestión de los digestores y las decisiones de despacho. En este contexto, resulta necesario evaluar hasta qué punto modelos de distinta complejidad permiten anticipar la producción diaria y capturar la estructura temporal de la serie.

Este trabajo analiza la predicción diaria de la generación eléctrica a partir de biogás en Galicia mediante una comparación entre seis enfoques: dos modelos de referencia, *naïve* y media móvil de siete días, un modelo estadístico SARIMA y tres arquitecturas recurrentes de aprendizaje profundo, RNN, LSTM y GRU. Para ello, se utilizaron registros reales procedentes del Sistema de Información del Operador del Sistema (ESIOS) de Red Eléctrica de España, correspondientes a A Coruña, Lugo, Orense y Pontevedra, agregados en una serie regional diaria para el periodo comprendido entre abril de 2016 y marzo de 2026. La evaluación se realizó con una partición cronológica de los datos y un horizonte de predicción de un día, empleando MAE y RMSE como métricas de desempeño en la escala original de generación. Los resultados muestran que la serie presenta persistencia diaria y una periodicidad semanal marcada, mientras que la estacionalidad mensual es poco relevante. En términos predictivos, la GRU obtuvo el menor error en el conjunto de prueba, con un MAE de 5,750 MWh y un RMSE de 7,448 MWh, lo que supone una mejora del 40,8 % frente al modelo *naïve*. Sin embargo, el modelo SARIMA alcanzó un desempeño competitivo, con un MAE de 6,136 MWh, apenas 0,386 MWh por encima de la GRU. Estos resultados indican que la principal mejora predictiva procede de representar adecuadamente la persistencia de corto plazo y el patrón semanal de la serie, más que del incremento de complejidad arquitectónica. En consecuencia, para esta serie univariante, modelos estadísticos bien especificados y arquitecturas recurrentes compactas ofrecen un equilibrio favorable entre precisión, interpretabilidad y complejidad metodológica.

Palabras clave: Biogás, predicción, series temporales, *Deep Learning* y Galicia.

Abstract

The forecasting of biogas-based electricity generation constitutes a relevant tool for the operational planning of anaerobic digestion plants and for the integration of this technology into power systems with a high share of renewable sources. Unlike resources such as wind or solar energy, the variability of biogas does not depend primarily on meteorological conditions, but rather on operational and biological factors linked to the type of substrate, the digestion conditions, the management of the digesters and dispatch decisions. In this context, it becomes necessary to assess to what extent models of varying complexity are able to anticipate daily production and capture the temporal structure of the series.

This study analyses the daily forecasting of biogas-based electricity generation in Galicia through a comparison among six approaches: two baseline models, a *naïve* method and a seven-day moving average, a statistical SARIMA model, and three recurrent deep learning architectures, namely RNN, LSTM and GRU. To this end, real records obtained from the Information System of the System Operator (ESIOS) of Red Eléctrica de España were used, corresponding to the provinces of A Coruña, Lugo, Orense and Pontevedra, aggregated into a single daily regional series for the period between April 2016 and March 2026. The evaluation was carried out using a chronological partition of the data and a one-day forecasting horizon, employing MAE and RMSE as performance metrics on the original generation scale. The results show that the series exhibits daily persistence and a marked weekly periodicity, whereas monthly seasonality proves to be of limited relevance. In predictive terms, the GRU achieved the lowest error on the test set, with an MAE of 5.750 MWh and an RMSE of 7.448 MWh, representing an improvement of 40.8 % over the *naïve* model. Nevertheless, the SARIMA model attained a competitive performance, with an MAE of 6.136 MWh, only 0.386 MWh above that of the GRU. These results indicate that the main predictive gain derives from adequately representing the short-term persistence and the weekly pattern of the series, rather than from increasing architectural complexity. Consequently, for this univariate series, well-specified statistical models and compact recurrent architectures offer a favourable balance between accuracy, interpretability and methodological complexity.

Keywords: Biogas, prediction, time series, Deep Learning, and Galicia.

Agradecimientos

Me gustaría agradecer a mi tutora, Jenny Alexandra Cifuentes Quintero, por su orientación, disponibilidad y criterio a lo largo de este trabajo.

También quiero expresar mi agradecimiento a Galicia, la tierra en la que he crecido y de la que proceden los datos analizados. Poder aplicar los conocimientos adquiridos para comprender mejor una pequeña parte de su potencial energético ha sido uno de los aspectos más gratificantes de este estudio.

Índice general

Resumen	III
Abstract	IV
1. Introducción	1
1.1. Motivación	1
1.2. Objetivos del Trabajo de Fin de Grado	4
1.2.1. Objetivo general	4
1.2.2. Objetivos específicos	4
1.3. Estructura del documento	4
2. Estado del arte sobre la predicción de la generación eléctrica a partir de biogás	6
3. Estrategias de predicción de series temporales usando Deep Learning	13
3.1. Redes Neuronales Recurrentes	14
3.2. Long Short-term Memory	16
3.3. Gated Recurrent Unit	18
3.4. Métricas de evaluación	20
3.5. Modelos de referencia para la evaluación del desempeño predictivo	21
4. Desarrollo experimental y resultados	24
4.1. Diseño metodológico	24
4.2. Análisis exploratorio de los datos	28
4.2.1. Caracterización general de la serie y tratamiento de valores atípicos	28
4.2.2. Distribución y patrones temporales de generación	31
4.2.3. Tendencia, estacionalidad y dependencia temporal	34
4.3. Modelado Predictivo	38
4.3.1. Modelos de referencia: naïve y media móvil	39
4.3.2. Modelo SARIMA	42
4.3.3. Modelo RNN	46
4.3.4. Modelo LSTM	49

4.3.5. Modelo GRU	52
4.3.6. Comparativa global de modelos	56
5. Conclusiones	59
Bibliografía	65

Índice de figuras

1.1. Proyección de la producción mundial de biogás por regiones en el periodo 2024–2030. Fuente: Elaboración propia adaptada de International Energy Agency (2025)	2
3.1. Representación de una RNN desplegada en el tiempo. Elaboración propia adaptada de Goodfellow, Bengio, y Courville (2016).	15
3.2. Esquema comparativo de la arquitectura de RNN frente LSTM. Elaboración propia adaptada de Olah (2015).	17
3.3. Esquema de una red GRU. Elaboración propia adaptada de Abdulwahab, Jabreel, y Moreno (2017).	19
4.1. Esquema metodológico del trabajo. Fuente: Elaboración propia.	25
4.2. Evolución diaria de la generación eléctrica a partir de biogás en Galicia. Fuente: Elaboración propia.	29
4.3. Identificación de valores atípicos sobre la serie temporal. Fuente: Elaboración propia.	30
4.4. Histograma y estimación de densidad por núcleo de la generación diaria de biogás en Galicia. Fuente: Elaboración propia.	31
4.5. Distribución de la generación diaria de electricidad a partir de biogás por mes del año. Fuente: Elaboración propia.	32
4.6. Distribución de la generación diaria de electricidad a partir de biogás por día de la semana. Fuente: Elaboración propia.	33
4.7. Variación mensual de la generación por año completo (2017 - 2025). Fuente: Elaboración propia.	34
4.8. Generación media diaria por año completo (2017 - 2025). Fuente: Elaboración propia.	34
4.9. Generación diaria y media móvil de 30 días. Fuente: Elaboración propia.	35
4.10. Descomposición estacional aditiva de la serie de generación. Fuente: Elaboración propia.	36
4.11. Función de autocorrelación y función de autocorrelación parcial de la serie. Fuente: Elaboración propia.	37

4.12. Generación mensual total en Galicia, MWh, en años completos. Fuente: Elaboración propia.	38
4.13. Comparación entre los valores reales y las predicciones de los modelos <i>naïve</i> y media móvil en el conjunto de prueba. Fuente: Elaboración propia.	41
4.14. Distribución del error absoluto del modelo <i>naïve</i> en entrenamiento, validación y prueba. Fuente: Elaboración propia.	42
4.15. Distribución del error absoluto del modelo de media móvil en entrenamiento, validación y prueba. Fuente: Elaboración propia.	42
4.16. Predicciones del modelo SARIMA $(1, 0, 0)(1, 0, 1)_7$ frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.	45
4.17. Evolución de la pérdida Huber durante el entrenamiento del modelo RNN final. Fuente: Elaboración propia.	47
4.18. Evolución del MAE en entrenamiento y validación durante el entrenamiento del modelo RNN final. Fuente: Elaboración propia.	48
4.19. Predicciones del modelo RNN frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.	49
4.20. Evolución de la pérdida Huber durante el entrenamiento del modelo LSTM final. Fuente: Elaboración propia.	51
4.21. Evolución del MAE en entrenamiento y validación durante el entrenamiento del modelo LSTM final. Fuente: Elaboración propia.	51
4.22. Predicciones del modelo LSTM frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.	52
4.23. Evolución de la pérdida Huber durante el entrenamiento del modelo GRU final. Fuente: Elaboración propia.	54
4.24. Evolución del MAE entrenamiento y validación durante el entrenamiento del modelo GRU final. Fuente: Elaboración propia.	54
4.25. Predicciones del modelo GRU frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.	55

Índice de tablas

2.1. Resumen de los estudios sobre la generación eléctrica a partir de biogás y de fuentes bioenergéticas similares.	12
4.1. Variables contenidas en los archivos CSV descargados de ESIOS.	26
4.2. Distribución de fechas según el número de provincias con dato registrado. .	26
4.3. Estructura de la base de datos de generación de biogás en Galicia.	28
4.4. Estadísticos descriptivos de la serie de generación de biogás, MWh/día. . .	29
4.5. Métricas de evaluación de los modelos de referencia. <i>naïve</i> y media móvil .	40
4.6. Top 5 configuraciones obtenidas para el modelo SARIMA mediante búsqueda en grilla.	44
4.7. Métricas de evaluación del modelo SARIMA $(1, 0, 0)(1, 0, 1)_7$	44
4.8. Top 5 configuraciones obtenidas para la RNN mediante <i>Random Search</i> . . .	47
4.9. Métricas finales del modelo RNN.	48
4.10. Top 5 configuraciones obtenidas para la LSTM mediante <i>Random Search</i> . .	50
4.11. Métricas finales del modelo LSTM.	51
4.12. Top 5 configuraciones obtenidas para la GRU mediante <i>Random Search</i> . . .	53
4.13. Métricas finales del modelo GRU.	55
4.14. Comparativa global de modelos según métricas MAE y RMSE (MWh). Fuente: Elaboración propia.	56

Acrónimos

ANN	Redes neuronales artificiales (<i>Artificial Neural Networks</i>)
ARIMA	Modelo autorregresivo integrado de medias móviles (<i>Autoregressive Integrated Moving Average</i>)
CH ₄	Metano
CNN	Red neuronal convolucional (<i>Convolutional Neural Network</i>)
CO ₂	Dióxido de carbono
DA	Atención dual (<i>Dual Attention</i>)
DL	Aprendizaje profundo (<i>Deep Learning</i>)
EDAR	Estación depuradora de aguas residuales
ESIOS	Sistema de Información del Operador del Sistema
GA	Algoritmo genético (<i>Genetic Algorithm</i>)
GRU	Unidades recurrentes cerradas (<i>Gated Recurrent Units</i>)
IEA	Agencia Internacional de la Energía (<i>International Energy Agency</i>)
IQR	Rango intercuartílico (<i>Interquartile Range</i>)
kWh	Kilovatio-hora
LSTM	Red neuronal de memoria a corto y largo plazo (<i>Long Short-Term Memory</i>)
MAE	Error absoluto medio (<i>Mean Absolute Error</i>)
MAPE	Error porcentual absoluto medio (<i>Mean Absolute Percentage Error</i>)
ML	Aprendizaje automático (<i>Machine Learning</i>)
MSE	Error cuadrático medio (<i>Mean Squared Error</i>)
MWh	Megavatios-hora
pH	Potencial de hidrógeno
PJ	Petajulio
R^2	Coefficiente de determinación
REE	Red Eléctrica de España
RMSE	Raíz del error cuadrático medio (<i>Root Mean Squared Error</i>)
RNN	Redes neuronales recurrentes (<i>Recurrent Neural Networks</i>)
SARIMA	Modelo autorregresivo integrado de medias móviles estacional (<i>Seasonal Autoregressive Integrated Moving Average</i>)

SEDIGAS	Asociación Española del Gas
TFT	Transformador de fusión temporal (<i>Temporal Fusion Transformer</i>)
VSN	Red de selección de variables (<i>Variable Selection Network</i>)

Capítulo 1

Introducción

1.1. Motivación

El biogás es un gas combustible que se genera cuando la materia orgánica se descompone en ausencia de oxígeno, mediante un proceso conocido como digestión anaerobia. Habitualmente procede de residuos agroganaderos, como purines y estiércoles, de la fracción orgánica de los residuos urbanos y de lodos de depuradora. Su composición está dominada por metano (CH_4), y dióxido de carbono (CO_2) (International Energy Agency, 2025). A partir del biogás bruto es posible obtener biometano mediante un proceso de depuración o *upgrading*, que elimina la mayor parte del CO_2 y otros compuestos no deseados. El gas resultante presenta propiedades similares a las del gas natural fósil, lo que permite su inyección en la red gaseísta o su uso directo como combustible en aplicaciones térmicas, eléctricas y de transporte (Ministerio para la Transición Ecológica y el Reto Demográfico, 2022). Por tanto, el biogás y el biometano forman parte de la bioenergía moderna y ofrecen una vía para transformar residuos en energía renovable almacenable y gestionable (International Renewable Energy Agency, 2025).

La transición hacia sistemas energéticos bajos en carbono constituye uno de los principales retos de las próximas décadas. Los informes recientes de la *International Energy Agency* (IEA), y del *Intergovernmental Panel on Climate Change*, coinciden en que el cumplimiento de los objetivos del Acuerdo de París 2015 exige reducir de forma sostenida las emisiones de gases de efecto invernadero, limitar el uso de combustibles fósiles y mantener la seguridad de suministro en un contexto geopolítico más incierto (IEA Bioenergy, 2022; IPCC, 2022). En este marco, los gases renovables han dejado de ocupar una posición marginal y se han incorporado a los escenarios de descarbonización y a las estrategias de diversificación energética (International Energy Agency, 2025; International Renewable Energy Agency, 2025).

A escala internacional, el biogás y el biometano permiten valorizar residuos agroganaderos, industriales y urbanos mediante su transformación en energía renovable. La digestión anaerobia evita la descomposición incontrolada de estos residuos, reduce emisiones de CH_4 y

contribuye a disminuir problemas de contaminación de suelos y aguas. Además, el digestato generado puede utilizarse como fertilizante bajo determinadas condiciones (IEA Bioenergy, 2022).

La relevancia creciente del biogás se refleja en las previsiones internacionales de producción. Como se muestra en la Figura 1.1, la producción mundial pasaría de aproximadamente 1.850 petajulios (PJ), anuales en 2024 a cerca de 2.350 PJ anuales en 2030, lo que representa un incremento de 500 PJ en el periodo analizado. La mayor contribución al crecimiento previsto procede de Europa, con un aumento aproximado de 230 a 250 PJ anuales. Esta aportación supera con claridad la estimada para otras regiones, como China, India, Estados Unidos y el resto de países considerados. La figura permite justificar la atención prestada al caso europeo, ya que muestra que el continente concentra la mayor parte del incremento esperado de producción de biogás hasta 2030.

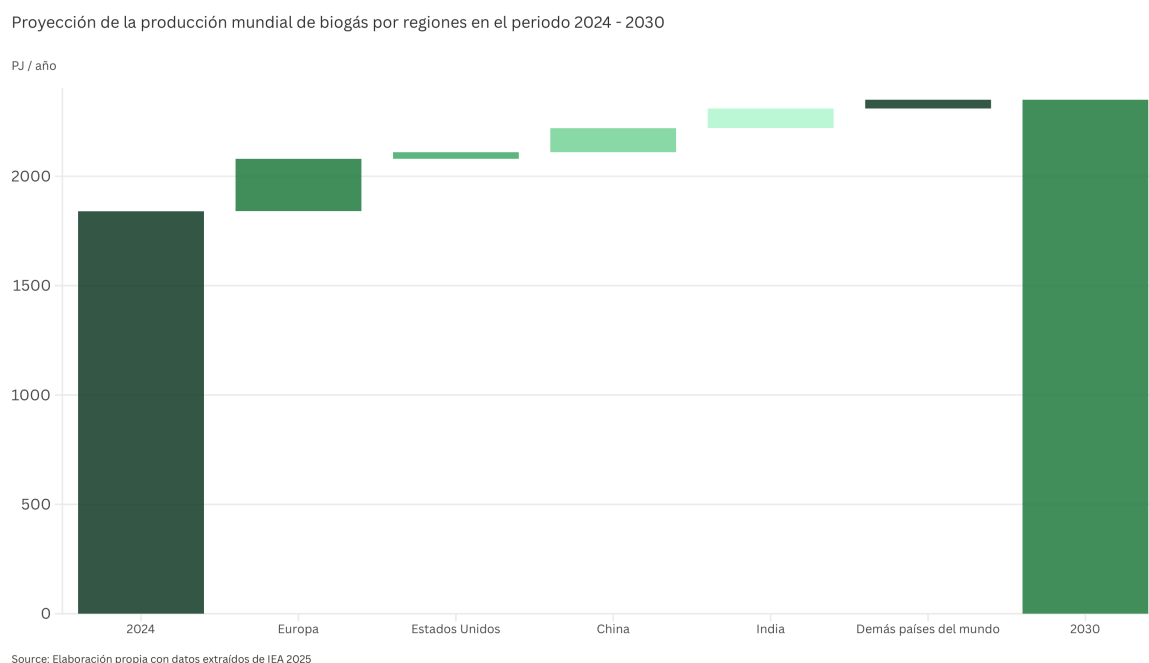


Figura 1.1: Proyección de la producción mundial de biogás por regiones en el periodo 2024–2030. Fuente: Elaboración propia adaptada de International Energy Agency (2025)

La estrategia REPowerEU responde a este escenario mediante una reducción acelerada de la dependencia de combustibles fósiles importados y establece como meta alcanzar una producción anual de 35 mil millones de metros cúbicos de biometano en 2030 (European Commission, 2022, 2024). Esta meta no se limita a sustituir parcialmente el gas natural fósil, sino que también promueve un uso más eficiente de residuos y subproductos generados dentro del territorio europeo. En este marco, el biogás y el biometano se consolidan como recursos energéticos compatibles con la descarbonización, la economía circular y la seguridad de suministro. Además, su carácter gestionable les permite complementar la generación

eólica y solar, aportando capacidad de respaldo en sistemas eléctricos con alta penetración de renovables variables (IEA Bioenergy, 2022).

España participa de forma directa en esta orientación europea. El Plan Nacional Integrado de Energía y Clima 2021–2030, y la Hoja de Ruta del Biogás reconocen el papel de los gases renovables en la reducción de la dependencia energética exterior, la disminución de emisiones y la valorización de residuos (Ministerio para la Transición Ecológica y el Reto Demográfico, 2022, 2023). La Hoja de Ruta establece un objetivo mínimo de 10,41 teravatios hora (TWh) anuales de producción de biogás en 2030 y señala que el potencial técnico asociado a residuos agroindustriales, biorresiduos municipales y lodos de depuradora supera ampliamente los niveles actuales de producción (Ministerio para la Transición Ecológica y el Reto Demográfico, 2022). En la misma línea, la Asociación Española del Gas (SEDIGAS), estima que el potencial accesible de producción de biometano podría alcanzar los 163 TWh anuales, equivalentes a aproximadamente el 43 % de la demanda nacional de gas natural registrada en 2021 (SEDIGAS, 2023). La distancia entre estas estimaciones y el grado de desarrollo actual del sector indica que el biogás continúa siendo un recurso poco explotado dentro del sistema energético español.

Dentro de este marco nacional, Galicia representa un caso de estudio pertinente por su estructura productiva y por la disponibilidad de residuos orgánicos procedentes del sector primario. La comunidad cuenta con una elevada densidad de explotaciones ganaderas, especialmente vinculadas a la producción bovina, lo que genera volúmenes relevantes de purines y estiércoles con potencial metanogénico. SEDIGAS sitúa a Galicia entre las comunidades autónomas con mayor potencial de producción de biometano, asociado principalmente al sector agroganadero y, en menor medida, a residuos forestales y urbanos (SEDIGAS, 2023). Sin embargo, la penetración efectiva del biogás en la potencia eléctrica renovable instalada de la región sigue siendo reducida en comparación con tecnologías consolidadas como la hidráulica y la eólica (Instituto Enerxético de Galicia, 2025). Esta diferencia entre disponibilidad de recursos y desarrollo energético efectivo muestra un margen de crecimiento para nuevos proyectos de biogás y biometano en el territorio gallego.

La relevancia del caso gallego se acentúa por la configuración de su sistema eléctrico, caracterizado por una presencia significativa de generación renovable variable. En este tipo de sistemas, el biogás puede contribuir a la cobertura de la demanda en momentos de menor producción eólica o solar, siempre que su operación se integre de forma adecuada en la planificación energética (IEA Bioenergy, 2022). Para ello, resulta necesario disponer de herramientas que permitan anticipar la generación esperada en horizontes de corto plazo. La predicción de la generación eléctrica procedente del biogás, basada en datos históricos y técnicas de series temporales, permite reducir la incertidumbre operativa, mejorar la programación de la producción y aportar información útil para decisiones de inversión y planificación territorial. Desde esta perspectiva, se plantea el presente Trabajo de Fin de Grado, centrado en el análisis y la predicción a corto plazo de la generación eléctrica a partir del biogás en

Galicia.

1.2. Objetivos del Trabajo de Fin de Grado

1.2.1. Objetivo general

El objetivo general de este Trabajo de Fin de Grado es formular y estimar un modelo de predicción a corto plazo de la energía eléctrica generada a partir del biogás en Galicia, empleando técnicas de análisis de series temporales y comprobando su capacidad predictiva a partir de datos históricos de generación.

1.2.2. Objetivos específicos

Para alcanzar este objetivo general se plantean los siguientes objetivos específicos:

- Contextualizar la importancia de disponer de predicciones a corto plazo de la generación eléctrica procedente del biogás en Galicia, analizando el papel de esta tecnología en el sistema energético regional y su contribución potencial a la planificación operativa y a la integración de energías renovables.
- Identificar y revisar las principales metodologías de predicción aplicables a series temporales de generación energética, valorando su idoneidad para el caso del biogás a partir de la literatura científica y técnica disponible.
- Diseñar y construir un modelo de predicción a corto plazo basado en los registros históricos de generación eléctrica atribuida al biogás en Galicia, incorporando aquellas técnicas que mejor se adapten al comportamiento observado de la serie.
- Evaluar y validar el modelo propuesto mediante la comparación de sus resultados con modelos de referencia (*baseline*) y con otras técnicas de predicción seleccionadas, utilizando métricas de error habituales en *forecasting* para cuantificar su desempeño.

1.3. Estructura del documento

La estructura de esta memoria se organiza en cinco capítulos. El Capítulo 2 ofrece un marco conceptual y una revisión de la literatura relacionada con el biogás y el biometano, su papel en los sistemas energéticos a escala internacional, nacional y regional, así como los principales enfoques de predicción aplicados a series temporales de generación energética. El Capítulo 3 presenta los fundamentos teóricos de las estrategias de predicción consideradas, con especial atención a los modelos de referencia y a las arquitecturas recurrentes de

aprendizaje profundo (RNN, LSTM y GRU), así como a las métricas empleadas para evaluar su desempeño. En el Capítulo 4 se detallan las fuentes de datos utilizadas y la metodología seguida: se especifican los criterios de selección de las series de generación eléctrica a partir de biogás en Galicia, las etapas de preprocesamiento y análisis descriptivo, y la formulación de los modelos de predicción considerados. A continuación, se presenta la implementación empírica de dichos modelos, evaluando su desempeño mediante las métricas de error seleccionadas y comparando los resultados obtenidos con los modelos *baseline*. Finalmente, el Capítulo 5 recoge las principales conclusiones del trabajo, discute sus implicaciones para la planificación energética en Galicia y plantea posibles líneas de investigación futura orientadas a profundizar y ampliar los resultados presentados.

Capítulo 2

Estado del arte sobre la predicción de la generación eléctrica a partir de biogás

La predicción a corto plazo de la generación eléctrica a partir de biogás constituye una herramienta relevante para la operación diaria de plantas de digestión anaerobia y para la gestión de sistemas eléctricos con una elevada participación de energías renovables. Las estimaciones de producción se utilizan para programar la generación, gestionar la inyección de energía en la red y optimizar el funcionamiento de los equipos de conversión energética. En el caso específico del biogás, disponer de previsiones fiables permite ajustar la alimentación de los digestores, anticipar cambios en el rendimiento energético y reducir ineficiencias asociadas a variaciones en la composición del sustrato o en las condiciones de operación (Rieke, 2018). Asimismo, la incorporación progresiva de fuentes renovables variables en los sistemas eléctricos ha incrementado la necesidad de modelos capaces de anticipar la disponibilidad energética en horizontes temporales reducidos. Las primeras aproximaciones empleadas para este tipo de predicción se basaron principalmente en modelos estadísticos lineales de series temporales. Entre los enfoques más utilizados destacan el modelo autorregresivo integrado de medias móviles *Autoregressive Integrated Moving Average* (ARIMA) y los métodos de suavizado exponencial como *Holt–Winters*. Estos modelos han sido ampliamente aplicados en el ámbito energético debido a su bajo coste computacional, su facilidad de implementación y la interpretabilidad de sus parámetros (Box, Jenkins, Reinsel, y Ljung, 2015). Su funcionamiento se apoya en la identificación de patrones temporales persistentes, como tendencias, estacionalidad y autocorrelación, presentes en las observaciones históricas.

En el ámbito del biogás, Şenol, Çolak, y Oda (2024) emplean modelos ARIMA para estimar el potencial regional de producción de biogás en Turquía y reportan errores porcentuales absolutos medios (*Mean Absolute Percentage Error*, MAPE) inferiores al 5 % en series agregadas y relativamente estables. Los autores concluyen que el modelo ofrece un comportamiento adecuado cuando las dinámicas temporales presentan regularidad y una volatilidad limitada. De forma similar, Ribeiro, Marques, Junior, de Almeida, y Campos (2019) aplican

Holt–Winters a la producción mensual de gas natural en Brasil y obtienen resultados satisfactorios en la captura de tendencias y patrones estacionales. Aunque el estudio no se centra específicamente en biogás, los resultados son relevantes debido a las similitudes existentes entre ambos contextos energéticos en términos de comportamiento temporal agregado y necesidades de planificación operativa. Estos trabajos muestran que los modelos lineales ofrecen un desempeño competitivo cuando las series presentan una estructura temporal relativamente regular y una dependencia moderada de variables externas. No obstante, la producción de biogás en condiciones reales depende de procesos biológicos complejos influenciados por factores como la composición del sustrato, la temperatura, el potencial de hidrógeno (pH), el tiempo de retención hidráulica y la actividad microbiológica (Dittmer et al., 2021). Como consecuencia, las series de producción suelen incorporar no linealidades, cambios abruptos de régimen y episodios de alta variabilidad que resultan difíciles de representar mediante enfoques estadísticos tradicionales.

La creciente disponibilidad de datos procedentes de sensores industriales y sistemas de monitorización ha favorecido la incorporación de técnicas *Deep Learning* (DL), orientadas al modelado de secuencias temporales complejas. Dentro de este contexto, las redes neuronales recurrentes (*Recurrent Neural Networks*, RNN) han adquirido una presencia significativa debido a su capacidad para incorporar dependencias temporales entre observaciones consecutivas (Lim y Zohren, 2021). Sin embargo, las RNN convencionales presentan limitaciones relacionadas con el desvanecimiento y explosión del gradiente durante el entrenamiento, especialmente en secuencias largas (Bengio, Simard, y Frasconi, 1994). Con el objetivo de resolver estas limitaciones surgieron arquitecturas recurrentes avanzadas como la red neuronal de memoria a corto y largo plazo (*Long Short–Term Memory*, LSTM) y las unidades recurrentes cerradas (*Gated Recurrent Units*, GRU). Ambas arquitecturas incorporan mecanismos de compuertas que permiten retener información relevante durante horizontes temporales prolongados y descartar señales menos útiles para la predicción. Las LSTM utilizan estructuras de memoria interna y múltiples compuertas de entrada, salida y olvido para controlar el flujo de información temporal (Hochreiter y Schmidhuber, 1997). Por su parte, las GRU simplifican parcialmente esta estructura mediante un menor número de compuertas y parámetros entrenables, lo que reduce el coste computacional y acelera el entrenamiento (Cho et al., 2014).

En aplicaciones energéticas, estas arquitecturas se han utilizado de forma extensa para predicción de demanda eléctrica, producción renovable y sistemas de almacenamiento energético (Lim y Zohren, 2021). Los estudios revisados coinciden en que las LSTM y GRU ofrecen ventajas frente a modelos lineales cuando la serie presenta relaciones no lineales complejas, dependencias temporales de largo alcance y elevada variabilidad operativa. No obstante, su desempeño depende de la disponibilidad de conjuntos de datos suficientemente amplios y representativos, además de un proceso cuidadoso de ajuste de hiperparámetros para evitar problemas de sobreajuste y pérdida de capacidad de generalización. En el ám-

bito específico del biogás, las arquitecturas LSTM y GRU han comenzado a aplicarse tanto a la predicción de producción de gas como a la estimación de generación eléctrica asociada. Oliveira et al. (2024) analizan una Estación Depuradora de Aguas Residuales (EDAR), en Portugal y desarrollan modelos basados en LSTM y GRU utilizando datos diarios de generación, variables meteorológicas y parámetros operativos. Sus resultados muestran que la arquitectura LSTM alcanza una raíz del error cuadrático medio (*Root Mean Squared Error*, RMSE) aproximado de 347,9 kilovatio-hora (kWh), superando tanto a la GRU como a modelos más simples utilizados como referencia. Los autores destacan que la incorporación simultánea de variables operativas y meteorológicas mejora la capacidad predictiva del modelo al capturar interacciones no lineales presentes en el proceso de digestión anaerobia. De manera complementaria, Jeong et al. (2021) proponen un modelo híbrido basado en mecanismos de atención dual y selección automática de variables para predecir la producción de biogás en procesos de codigestión anaerobia de residuos orgánicos. Su arquitectura combina una LSTM con un mecanismo *Dual Attention* (DA) y una red de selección de variables (*Variable Selection Network*, VSN). El estudio utiliza dos años de datos reales de monitorización y muestra mejoras sustanciales respecto a una LSTM convencional. En concreto, el coeficiente de determinación (R^2) aumenta desde 0,38 en la LSTM estándar hasta 0,68 en la arquitectura DA-LSTM y 0,76 en el modelo DA-LSTM-VSN. Además de mejorar la precisión predictiva, el enfoque permite identificar qué variables de entrada presentan una mayor influencia sobre la producción de biogás, aportando un grado adicional de interpretabilidad al modelo.

Sobre esta base, Salamattalab, Hasani Zonoozi, y Molavi-Arabshahi (2024) proponen un enfoque orientado a mejorar la selección de variables en digestores anaerobios a gran escala. El estudio se desarrolla en una EDAR de Teherán y emplea un modelo híbrido basado en un *Genetic Algorithm* (GA), y LSTM para predecir la producción diaria de biogás a partir de datos de aguas residuales y corrientes de fango. La ventana temporal de entrada se fija en 21 días, en correspondencia con el tiempo de retención hidráulica de los digestores analizados. El GA permite seleccionar subconjuntos reducidos de variables de entrada, mientras que la LSTM modela las dependencias temporales de la serie. Los resultados muestran valores de R^2 entre 0,84 y 0,90 en tres escenarios de modelización, lo que indica un rendimiento predictivo elevado con un número limitado de parámetros. Este tipo de enfoque resulta relevante porque aborda simultáneamente dos problemas habituales en la predicción de biogás, la elevada dimensionalidad de las variables operativas y la necesidad de capturar relaciones temporales no lineales.

No obstante, tanto en este trabajo como en los estudios anteriores, la dinámica temporal se trata principalmente mediante bloques recurrentes. Esta aproximación puede ser insuficiente cuando las series presentan ruido, fluctuaciones de corta duración o cambios operativos rápidos. En estos casos, la incorporación de módulos adicionales capaces de extraer patrones locales antes de la etapa recurrente puede mejorar la representación de la señal temporal. En

el ámbito de la demanda eléctrica, Yildiz Dogan, Aksoy, y Öztürk (2024) muestran que un modelo híbrido basado en un *Convolutional Neural Network* (CNN), y LSTM, aplicado a series horarias de consumo eléctrico y variables meteorológicas en una ciudad de Turquía, alcanza un R^2 de 0,8599, superior al R^2 de 0,8086 obtenido por una LSTM estándar. Además, el modelo CNN–LSTM reduce de forma consistente el error porcentual absoluto medio (*Mean Absolute Percentage Error*, MAPE), el error absoluto medio (*Mean Absolute Error*, MAE), y el RMSE. Estos resultados sugieren que la etapa convolucional contribuye a identificar variaciones locales y cambios rápidos en la carga, lo que mejora la robustez del modelo frente a fluctuaciones de corto plazo.

Las aplicaciones directas de arquitecturas CNN–LSTM al biogás siguen siendo limitadas, pero la estructura del problema presenta similitudes con otros contextos energéticos. Las series de generación eléctrica a partir de biogás combinan patrones de fondo asociados al funcionamiento general del sistema con variaciones rápidas derivadas de la operación diaria, cambios en la carga de sustrato, incidencias técnicas o alteraciones en las condiciones del digester. Por esta razón, los modelos híbridos CNN–LSTM resultan metodológicamente pertinentes, ya que permiten separar la extracción de patrones locales, a cargo de la CNN, del modelado de dependencias temporales, a cargo de la LSTM. Esta línea de trabajo conduce también a arquitecturas multiescala, en las que la señal se descompone en distintos niveles de frecuencia antes de ser procesada por el modelo predictivo.

Una forma de avanzar en esta dirección consiste en incorporar técnicas de descomposición multiescala, como la *Discrete Wavelet Transform*. Bajo este enfoque, conocido como Wavelet–CNN–LSTM, la serie temporal se descompone inicialmente en componentes de baja y alta frecuencia, que posteriormente son procesadas mediante capas convolucionales y recurrentes (Chen, Ma, Li, Wang, y Li, 2023). Scopinho y Buratto (2024) desarrollan esta aproximación en series temporales de generación eléctrica procedente de sistemas térmicos de biomasa. En su estudio, la señal se descompone mediante transformada *wavelet* y después se procesa con una arquitectura CNN–LSTM, lo que permite representar de forma separada la tendencia, las fluctuaciones de alta frecuencia y el ruido. El modelo híbrido obtiene un MAPE de 0,0148, equivalente aproximadamente al 1,5 %, y supera en precisión a otros modelos de DL evaluados en el trabajo. Estos resultados muestran el potencial de las arquitecturas multiescala para mejorar la predicción de generación eléctrica cuando la serie combina componentes temporales de distinta naturaleza. El caso de estudio se centra en biomasa sólida, pero la lógica metodológica resulta trasladable a la generación eléctrica a partir de biogás. En ambos contextos coexisten factores de largo plazo, como el diseño del sistema, la capacidad instalada y el tipo de recurso energético utilizado, con perturbaciones de corto plazo asociadas a cambios de carga, variaciones operativas o incidencias en el proceso. Pese a ello, estos modelos presentan limitaciones relevantes. Su diseño es más complejo que el de las arquitecturas recurrentes simples, requieren ajustar los parámetros de la descomposición *wavelet* y suelen implicar un mayor coste computacional. Además, la evidencia empírica disponible

en plantas de biogás todavía es reducida, por lo que su aplicación requiere una evaluación cuidadosa en función del tamaño, la granularidad y la calidad de los datos disponibles.

En paralelo, la literatura reciente ha explorado modelos basados en mecanismos de atención, especialmente la familia Transformer, como alternativa a las estructuras recurrentes tradicionales. Estos modelos permiten capturar dependencias de largo alcance y manejar múltiples variables explicativas sin depender necesariamente de una transmisión secuencial de la información (Chen et al., 2023; Wen et al., 2023). Su aplicación al biogás es todavía incipiente, pero los primeros resultados disponibles muestran un desempeño prometedor. Sappl, Harders, y Rauch (2023) calibran un *Temporal Fusion Transformer* (TFT), a partir de seis años de datos de operación a escala real en digestores anaerobios de una EDAR. El modelo combina series temporales de tasas de producción de biogás con variables categóricas, como días festivos, y genera predicciones cuantílicas para los siete días siguientes. En comparación con tres técnicas estadísticas de referencia, el TFT alcanza valores de MAPE inferiores al 8 % y proporciona intervalos de predicción más robustos frente a las fluctuaciones de la serie. Además, la arquitectura permite obtener pesos de importancia de las variables, lo que aporta información explícita sobre los factores operativos y ambientales que contribuyen a cada predicción (Sappl et al., 2023). Estos resultados sitúan a los modelos basados en atención y, en particular, al TFT como una de las líneas más avanzadas en la predicción energética, ya que permiten integrar un número elevado de variables explicativas, capturar dependencias de largo alcance y generar salidas probabilísticas interpretables. Su adopción práctica exige, sin embargo, series históricas extensas y de buena calidad, recursos de cómputo suficientes y una fase de diseño más cuidadosa que la requerida por modelos estadísticos o redes recurrentes convencionales.

La Tabla 2.1 sintetiza los principales estudios identificados en la literatura reciente sobre predicción de generación energética a partir de biogás y de fuentes bioenergéticas análogas. La evidencia revisada muestra una transición desde modelos estadísticos clásicos, como ARIMA y Holt–Winters, hacia enfoques de DL y arquitecturas híbridas de mayor complejidad. Esta evolución no implica una sustitución directa de los modelos simples por modelos más sofisticados, ya que la elección metodológica depende del objetivo de predicción, del horizonte temporal, de la granularidad de los datos, del número de variables disponibles y del grado de variabilidad de la serie analizada. Los modelos lineales resultan adecuados como referencia inicial cuando las series son agregadas, relativamente estables y con patrones temporales regulares. Sin embargo, en contextos donde la generación depende de procesos biológicos, condiciones operativas cambiantes y relaciones no lineales, la literatura reporta mejores resultados con arquitecturas basadas en RNN, especialmente LSTM y GRU. Estas técnicas han mostrado un desempeño favorable tanto en sus versiones estándar como en variantes híbridas que incorporan mecanismos de selección de variables, capas convolucionales o descomposición multiescala. Por esta razón, su inclusión en el presente trabajo queda justificada como una alternativa metodológica adecuada para modelar la variabilidad temporal

propia de la generación eléctrica a partir de biogás.

El presente Trabajo de Fin de Grado se sitúa en esta línea de investigación y aborda la predicción de la generación eléctrica a partir de biogás desde una perspectiva comparativa. A partir de datos históricos, se evaluará el desempeño de modelos estadísticos clásicos como punto de referencia y de modelos basados RNN, en particular LSTM y GRU, como alternativas para capturar dependencias temporales y patrones no lineales en la serie.

Tabla 2.1: Resumen de los estudios sobre la generación eléctrica a partir de biogás y de fuentes bioenergéticas similares.

Referencia	Objetivo del Estudio	Tipo de Datos	Métodos Utilizados	Métricas de Evaluación	Principales Resultados
Şenol et al. (2024)	Estimar el potencial regional de producción de biogás y biometano en Turquía del 2022 al 2035.	Series anuales de población, censo ganadero, residuos agroganaderos y variables socioeconómicas (2004–2035).	ARIMA, redes neuronales artificiales (ANN) y modelos de regresión lineal.	MAPE.	ARIMA y ANN alcanzan MAPE inferiores al 5 %, mostrando buen desempeño cuando las series están agregadas y presentan patrones relativamente suaves.
Ribeiro et al. (2019)	Predecir la producción mensual de gas natural en Brasil y su estacionalidad.	Producción mensual de gas natural (2000–2017).	Holt–Winters aditivo y multiplicativo.	MAPE, MAE, error cuadrático medio (MSE).	El modelo Holt–Winters multiplicativo presenta un buen desempeño, mostrando un ajuste adecuado y predicciones mayoritariamente contenidas dentro del intervalo de confianza del 95 %.
Oliveira et al. (2024)	Evaluar LSTM y GRU para predecir la generación diaria de electricidad a partir de biogás en una EDAR.	Datos diarios de generación, meteorología y operación en una EDAR en Portugal.	LSTM, GRU.	RMSE.	LSTM reduce de forma significativa el error, RMSE en torno a 347,9 kWh, mejorando el rendimiento frente a GRU.
Jeong et al. (2021)	Desarrollar un modelo de DL para predecir la producción de biogás en codigestión anaerobia y analizar la importancia de las variables de entrada.	Datos de monitorización de dos años en un proceso de codigestión anaerobia en una EDAR.	LSTM estándar, DA–LSTM, modelo híbrido DA–LSTM–VSN.	R^2 .	La incorporación de DA y VSN aumenta R^2 , de un R^2 de 0,38 en LSTM a 0,68 en DA–LSTM y 0,76 en DA–LSTM–VSN), permitiendo interpretar la relevancia de las variables asociadas al lixiviado de residuos alimentarios.
Salamattalab et al. (2024)	Predecir la producción diaria de biogás y seleccionar variables óptimas en un digestor industrial.	Histórico diario de producción y 11 variables operativas en un digestor industrial de Irán.	LSTM + GA.	R^2 .	Se alcanzan valores de R^2 hasta 0,90 con una ventana de 21 días coherente con el tiempo de retención hidráulica.
Yildiz Dogan et al. (2024)	Proponer y evaluar un modelo híbrido CNN–LSTM para estimar la demanda eléctrica futura en ciudades sostenibles.	Series horarias de consumo eléctrico y variables meteorológicas de una ciudad de Turquía (2017–2021).	LSTM, modelo híbrido CNN–LSTM.	MAPE, RMSE, MAE, R^2 .	El modelo CNN–LSTM obtiene valores de MAPE, MAE y RMSE y un R^2 de 0,8599 frente a R^2 de 0,8086 del LSTM, mejorando la precisión del <i>forecasting</i> de carga eléctrica.
Scopinho y Buratto (2024)	Evaluar un modelo Wavelet–CNN–LSTM para predecir la generación eléctrica en sistemas térmicos de biomasa.	Serie temporal de la generación eléctrica de 3 meses en sistemas térmicos de biomasa en Brasil.	Wavelet–CNN–LSTM.	MAPE.	El modelo Wavelet–CNN–LSTM alcanza un MAPE de 0,0148 y supera en precisión de los modelos por separado.
Sapli et al. (2023)	Realizar regresión cuantílica de la tasa de producción de biogás en digestores anaerobios y comparar un modelo TFT con técnicas estadísticas clásicas.	Serie temporal multivariante de tasas de producción de biogás en una EDAR; seis años de registro.	TFT con regresión cuantílica; tres modelos estadísticos de referencia.	MAPE, MAE, RMSE.	El TFT alcanza MAPE <8 % en predicciones a siete días vista y supera a los modelos estadísticos de referencia, además de proporcionar cuantiles e importancia de variables que mejoran la interpretabilidad de las predicciones, $R^2 > 0,90$.

Capítulo 3

Estrategias de predicción de series temporales usando Deep Learning

En este capítulo se presentan los fundamentos teóricos de los modelos que servirán de base para el desarrollo empírico del trabajo. El objetivo es introducir las principales estrategias de predicción de series temporales consideradas en el estudio, con especial atención a los modelos basados en redes neuronales recurrentes y a su aplicación en problemas donde la información observada en un instante depende parcialmente de valores anteriores. Esta perspectiva resulta adecuada para la predicción de generación energética, ya que la producción registrada en cada periodo suele estar condicionada por la dinámica previa del sistema, las condiciones operativas y la persistencia temporal de la propia serie.

El aprendizaje profundo es una rama del *Machine Learning* (ML) basada en redes neuronales artificiales con múltiples capas (Goodfellow et al., 2016). Estos modelos permiten aproximar relaciones no lineales complejas a partir de los datos y aprender representaciones internas útiles para la tarea predictiva. En el contexto de las series temporales, su interés radica en la capacidad para modelar dependencias dinámicas, patrones no lineales y cambios de régimen que pueden resultar difíciles de capturar mediante especificaciones lineales tradicionales (Lim y Zohren, 2021). No obstante, su mayor flexibilidad también exige un diseño cuidadoso del entrenamiento, una selección adecuada de hiperparámetros y una evaluación rigurosa frente a modelos de referencia más simples (Hyndman y Athanasopoulos, 2021).

Por esta razón, el análisis no se limitará a modelos de aprendizaje profundo, también se incorporarán modelos de referencia que permitan valorar si el aumento de complejidad metodológica se traduce en una mejora predictiva real. En primer lugar, se considerará un modelo *naïve*, entendido como una estrategia de referencia simple basada en la persistencia o agregación de valores históricos recientes. Su especificación concreta podrá adoptar distintas formas, como utilizar el valor observado en un periodo previo comparable o el promedio de una ventana temporal anterior, según la granularidad y estructura de la serie. Adicionalmente, se incorporará un modelo basado en una media móvil de siete días. Esta aproximación

genera las predicciones a partir del promedio de las observaciones más recientes, suavizando fluctuaciones puntuales y capturando de forma sencilla patrones recurrentes de corto plazo. La elección de una ventana semanal resulta coherente con la periodicidad observada en la serie y permite disponer de una referencia adicional frente a la que evaluar la capacidad de los modelos más avanzados para extraer información temporal relevante. En segundo lugar, se ajustará un modelo estadístico tradicional basado en estructuras autorregresivas, en particular un modelo autorregresivo integrado de medias móviles estacional (*Seasonal Autoregressive Integrated Moving Average*, SARIMA). Este modelo permite representar dependencia temporal, tendencia y patrones estacionales mediante una formulación paramétrica interpretable. La comparación de enfoques permitirá evaluar si las arquitecturas basadas en redes neuronales aportan una mejora predictiva frente a referencias simples y estadísticamente consolidadas.

A partir de esta base, el capítulo se organiza en torno a las arquitecturas que serán consideradas en el estudio empírico. En primer lugar, se introducen las RNN, como punto de partida para el modelado de datos secuenciales. Posteriormente, se describen las arquitecturas LSTM, y las GRU, diseñadas para mejorar la conservación de información temporal relevante. Finalmente, se presentan las métricas de evaluación y junto con los modelos de referencia que permitirán comparar el desempeño de las distintas alternativas predictivas.

3.1. Redes Neuronales Recurrentes

Las RNN, constituyen una de las primeras arquitecturas diseñadas para el tratamiento de datos secuenciales. A diferencia de una red neuronal *feedforward*, en la que cada observación se procesa de manera independiente, una RNN incorpora conexiones recurrentes que permiten transferir información de un instante temporal al siguiente. De este modo, la predicción realizada en un momento determinado no depende únicamente de la entrada actual, sino también de una representación interna construida a partir de observaciones anteriores (Elman, 1990). Esta representación interna recibe el nombre de estado oculto. En cada instante t , la red actualiza dicho estado combinando la entrada observada x_t con el estado oculto procedente del instante anterior h_{t-1} . De forma general, esta actualización puede expresarse como:

$$h_t = \phi(W_x x_t + W_h h_{t-1} + b) \quad (3.1)$$

donde x_t representa la entrada en el instante t , h_{t-1} es el estado oculto anterior, W_x es la matriz de pesos asociada a la entrada, W_h es la matriz de pesos recurrente, b es el término de sesgo y ϕ corresponde a una función de activación no lineal, habitualmente una tangente hiperbólica o una función sigmoide. A partir de esta formulación, el nuevo estado h_t puede interpretarse como una síntesis de la información reciente y del historial acumulado por la red. Este estado puede utilizarse directamente para generar una salida o transmitirse al siguiente

instante temporal para continuar el procesamiento de la secuencia.

La Figura 3.1 representa una RNN desplegada en el tiempo. Cada nodo corresponde al estado oculto de la red en un instante concreto. Las flechas horizontales muestran la transmisión de información entre estados consecutivos, de modo que h_t se construye a partir de h_{t-1} y condiciona, a su vez, el cálculo de h_{t+1} . Esta representación permite visualizar la idea central de la recurrencia, dado que el modelo conserva una memoria dinámica de la secuencia y actualiza dicha memoria conforme recibe nuevas observaciones. En problemas de series temporales, esta estructura permite que la red incorpore información sobre la evolución previa de la variable antes de generar una predicción futura.

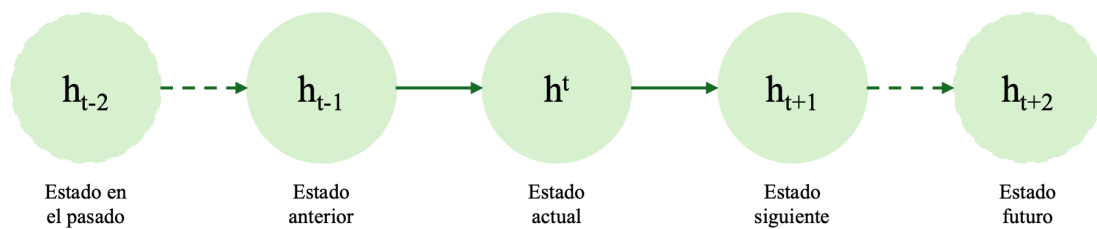


Figura 3.1: Representación de una RNN desplegada en el tiempo. Elaboración propia adaptada de Goodfellow et al. (2016).

No obstante, las RNN simples presentan limitaciones cuando la predicción requiere conservar información durante horizontes temporales prolongados. Durante el entrenamiento, el algoritmo *Backpropagation Through Time* propaga el error desde los últimos pasos de la secuencia hacia los primeros, actualizando los pesos de la red en función de las dependencias temporales acumuladas. En este proceso, los gradientes pueden reducirse progresivamente hasta volverse casi nulos, fenómeno conocido como desvanecimiento del gradiente, o crecer de forma inestable, dando lugar a la explosión del gradiente. Ambas situaciones dificultan la capacidad de las RNN convencionales para capturar dependencias de largo plazo (Bengio et al., 1994). En el contexto de la predicción de generación eléctrica a partir de biogás, esta limitación resulta especialmente relevante, dado que el comportamiento de la serie puede estar influido por procesos que no se manifiestan de forma inmediata, como cambios en la carga de sustrato, variaciones en el tiempo de retención hidráulica o alteraciones progresivas en las condiciones del digestor. Por este motivo, aunque las RNN ofrecen una base conceptual adecuada para el modelado de datos secuenciales, su uso suele ser sustituido o complementado por arquitecturas recurrentes más avanzadas, como LSTM y GRU, diseñadas para conservar información temporal durante periodos más prolongados.

3.2. Long Short-term Memory

Las redes neuronales LSTM, fueron propuestas por Hochreiter y Schmidhuber (1997) como una extensión de las redes neuronales recurrentes convencionales para superar las dificultades asociadas al aprendizaje de dependencias temporales de largo alcance. Su desarrollo surge como respuesta a las limitaciones observadas en las RNN simples, especialmente al problema del desvanecimiento y explosión del gradiente durante el entrenamiento, que reduce la capacidad de la red para transmitir información relevante a través de secuencias extensas. La principal diferencia entre una LSTM y una RNN convencional reside en la incorporación de una estructura de memoria explícita diseñada para almacenar y actualizar información durante horizontes temporales prolongados. Mientras una RNN actualiza un único estado oculto en cada instante, una LSTM introduce un estado de celda adicional que actúa como memoria interna del sistema. El acceso a esta memoria y su actualización se realizan mediante un conjunto de compuertas que regulan qué información debe conservarse, qué información nueva debe incorporarse y qué contenido almacenado debe utilizarse para generar la salida del modelo. El funcionamiento matemático de una unidad LSTM puede representarse mediante las siguientes ecuaciones:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3.2)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (3.3)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (3.4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (3.5)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (3.6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (3.7)$$

donde f_t representa la compuerta de olvido, i_t la compuerta de entrada, o_t la compuerta de salida, c_t el estado de celda y h_t el estado oculto. Los términos W y U corresponden a matrices de pesos asociadas a las entradas y estados ocultos previos, mientras que b representa el término de sesgo. La función sigmoide σ genera valores entre 0 y 1, permitiendo interpretar las compuertas como mecanismos de regulación del flujo de información. El operador $\odot$ representa el producto elemento a elemento. La secuencia de ecuaciones describe un proceso de actualización progresiva de la memoria interna. La compuerta de olvido determina qué proporción de la memoria anterior c_{t-1} debe mantenerse. Posteriormente, la compuerta de entrada controla qué información nueva candidata $\tilde{c}_t$ se incorpora al estado de celda. Finalmente, la compuerta de salida regula qué parte de la memoria acumulada se utiliza para generar el estado oculto que será transferido al siguiente instante temporal.

La Figura 3.2 muestra una comparación entre una RNN convencional y una arquitectura

LSTM desplegadas en el tiempo. En la parte superior se observa que la RNN transmite la información a través de una única trayectoria recurrente, donde el estado oculto se actualiza directamente a partir de la entrada y del estado previo. Esta estructura obliga a que toda la información relevante se concentre en una única representación interna. En la parte inferior se aprecia que la arquitectura LSTM incorpora un flujo adicional asociado al estado de celda y varios mecanismos de control representados mediante compuertas. Este diseño introduce una trayectoria de información más estable que permite conservar determinadas señales a lo largo de múltiples pasos temporales y reducir la pérdida progresiva de información durante el entrenamiento.

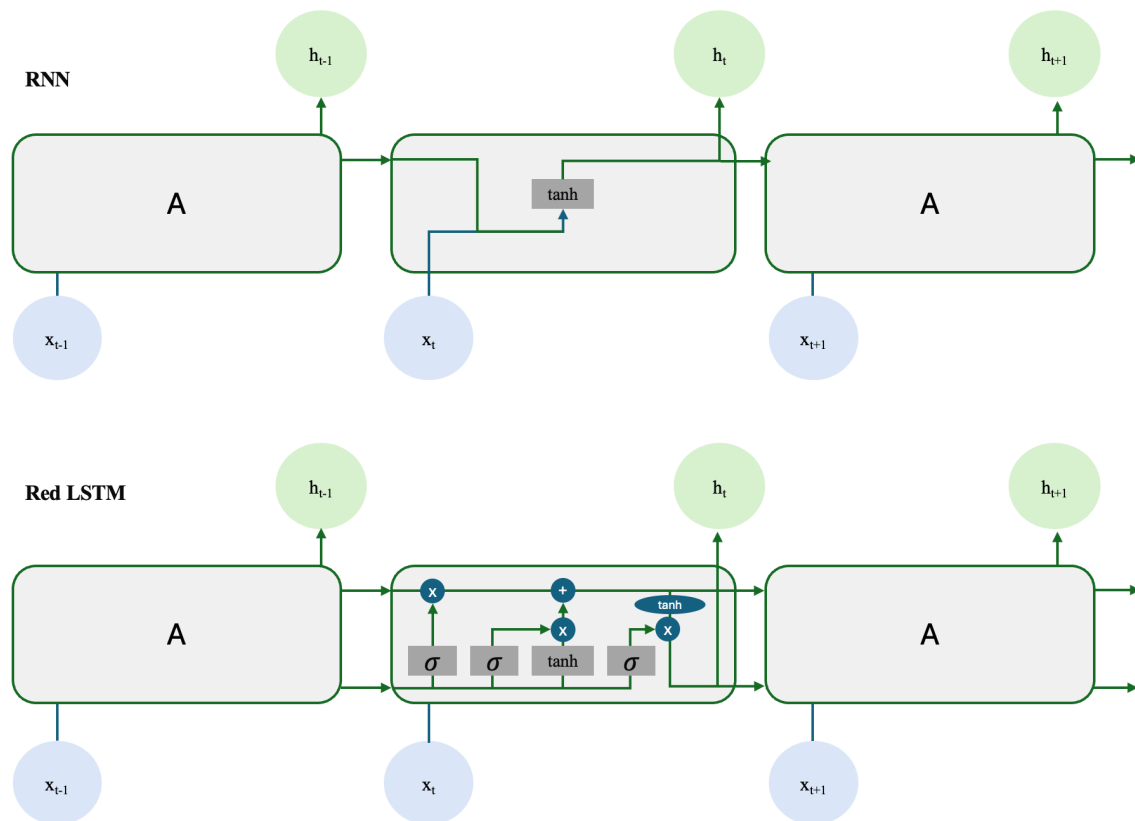


Figura 3.2: Esquema comparativo de la arquitectura de RNN frente LSTM. Elaboración propia adaptada de Olah (2015).

Desde la perspectiva de las series temporales, esta propiedad resulta particularmente adecuada cuando la predicción depende simultáneamente de efectos inmediatos y de patrones observados en horizontes temporales más amplios. En problemas de generación energética, el valor observado en un instante puede estar condicionado tanto por variaciones recientes como por comportamientos acumulados durante periodos anteriores. En el caso de la generación eléctrica a partir de biogás, la dinámica del sistema puede depender de procesos que evolucionan de manera gradual, como cambios en la composición del sustrato, variaciones en las condiciones de digestión o modificaciones en la carga de alimentación. Asimismo,

la incorporación de mecanismos adicionales de memoria implica también un aumento en la complejidad del modelo. Una LSTM contiene un número superior de parámetros entrenables respecto a una RNN convencional debido a la presencia de múltiples compuertas y estados internos (Hochreiter y Schmidhuber, 1997). Como consecuencia, el entrenamiento requiere un mayor coste computacional y puede presentar una mayor sensibilidad a la selección de hiperparámetros, al tamaño del conjunto de entrenamiento y a la calidad del preprocesamiento aplicado a los datos (Greff, Srivastava, Koutník, Steunebrink, y Schmidhuber, 2017). Por ello, aunque las LSTMs ofrecen una representación temporal más flexible que las RNNs simples, su desempeño depende de las características específicas del problema y deben evaluarse mediante comparación con otras alternativas de modelado.

3.3. Gated Recurrent Unit

Las redes neuronales GRU, fueron propuestas por Cho et al. (2014) como una arquitectura recurrente con compuertas orientada a modelar dependencias temporales de manera más eficiente que una RNN convencional. Al igual que las LSTM, las GRU regulan el flujo de información mediante mecanismos internos de control. Sin embargo, lo hacen con una estructura más compacta, ya que no separan explícitamente el estado de celda y el estado oculto. Esta simplificación reduce el número de parámetros entrenables y puede disminuir el coste computacional del entrenamiento. El funcionamiento de una GRU se basa principalmente en dos compuertas. La primera es la compuerta de actualización, encargada de determinar el grado en el cual el estado anterior se mantiene o se reemplaza por nueva información. La segunda es la compuerta de reinicio, responsable de regular la proporción del estado previo utilizada en la construcción del nuevo estado candidato. Su formulación habitual puede expresarse como:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (3.8)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (3.9)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \quad (3.10)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (3.11)$$

donde z_t representa la compuerta de actualización, r_t la compuerta de reinicio, $\tilde{h}_t$ el estado candidato y h_t el nuevo estado oculto. Los términos W y U corresponden a matrices de pesos asociadas a la entrada y al estado oculto previo, respectivamente, mientras que b representa el sesgo. La función sigmoide σ genera valores entre 0 y 1, por lo que las compuertas pueden interpretarse como filtros que regulan el paso de información. El operador $\odot$ denota el producto elemento a elemento.

La Figura 3.3 muestra el esquema interno de una unidad GRU. En ella se observa que el

estado previo h_{t-1} y la entrada actual x_t alimentan simultáneamente las compuertas de reinicio r_t y actualización z_t . La compuerta de reinicio interviene en la construcción del estado candidato $\tilde{h}_t$, al modular la influencia del pasado antes de aplicar la función $\tanh$. Posteriormente, la compuerta de actualización controla la combinación entre el estado anterior y el estado candidato para obtener h_t . De este modo, la GRU regula la proporción de memoria previa que se conserva y la proporción de información nueva que se incorpora en cada instante temporal.

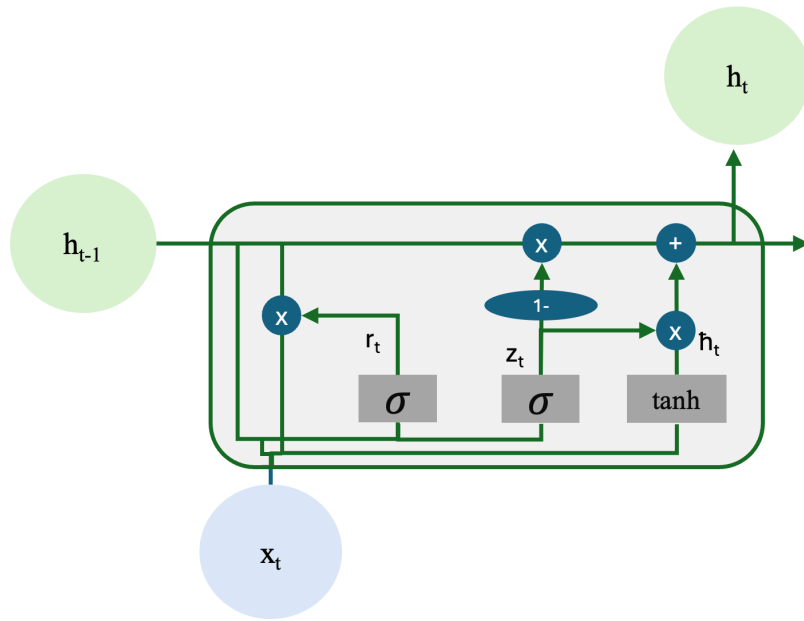


Figura 3.3: Esquema de una red GRU. Elaboración propia adaptada de Abdulwahab et al. (2017).

La principal diferencia respecto a una LSTM es que la GRU integra la memoria y la salida en un único estado oculto. Mientras que la LSTM dispone de un estado de celda separado para almacenar información de largo plazo, la GRU utiliza una representación más compacta en la que las compuertas controlan directamente la actualización del estado oculto. Esta característica puede resultar ventajosa cuando se trabaja con conjuntos de datos de tamaño moderado o cuando se busca reducir el tiempo de entrenamiento sin renunciar a la capacidad de capturar dependencias temporales (Chung, Gulcehre, Cho, y Bengio, 2014). Desde el punto de vista de la predicción de series temporales, las GRU ofrecen una alternativa intermedia entre la simplicidad de una RNN convencional y la mayor complejidad de una LSTM. Su menor número de parámetros puede favorecer un entrenamiento más estable en contextos donde la serie disponible no es extensa o donde el riesgo de sobreajuste debe controlarse con especial atención. En problemas de generación energética, esta propiedad resulta pertinente cuando se desea modelar dependencia temporal y no linealidad, pero se dispone de un volumen limitado de observaciones.

La comparación entre LSTM y GRU no debe plantearse como una jerarquía fija. Ambas arquitecturas incorporan compuertas y han mostrado un rendimiento superior al de las RNN simples en múltiples tareas secuenciales. La LSTM ofrece una estructura de memoria más diferenciada, lo que puede ser útil cuando existen dependencias largas y patrones temporales complejos. La GRU, por su parte, reduce la complejidad interna y puede alcanzar un desempeño comparable con menor coste computacional (Cho et al., 2014; Hochreiter y Schmidhuber, 1997). Por ello, en un estudio empírico de predicción resulta apropiado contrastar ambas arquitecturas bajo las mismas condiciones de datos, horizonte temporal y métricas de evaluación, con el fin de identificar la arquitectura que presenta mejor ajuste a la serie de generación eléctrica a partir de biogás.

3.4. Métricas de evaluación

Una vez estimados los modelos, la evaluación requiere criterios homogéneos para comparar su desempeño predictivo. En series temporales, la valoración debe realizarse sobre observaciones no utilizadas durante el entrenamiento, ya que el objetivo no consiste únicamente en ajustar la serie histórica, sino en generar predicciones fiables fuera de muestra. Por este motivo, la comparación entre modelos se basará en métricas de error calculadas sobre el conjunto de prueba. Entre las medidas más utilizadas se encuentran el RMSE, el MAE y el R^2 (Hyndman y Koehler, 2006).

El RMSE se define como:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.12)$$

donde y_i representa el valor observado, $\hat{y}_i$ el valor predicho y n el número de observaciones evaluadas. Esta métrica mide la magnitud promedio del error tras elevar al cuadrado las diferencias entre valores reales y estimados. Esta operación penaliza con mayor intensidad los errores grandes, por lo que el RMSE resulta sensible a desviaciones puntuales elevadas. Su interpretación se realiza en las mismas unidades de la variable objetivo y valores más bajos indican mejor desempeño predictivo (Hyndman y Koehler, 2006).

El MAE se expresa como:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.13)$$

Esta métrica calcula el promedio de las diferencias absolutas entre valores observados y predichos. A diferencia del RMSE, no amplifica los errores mediante una transformación cuadrática, por lo que ofrece una medida más estable ante valores extremos. Al expresar-

se también en las unidades originales de la variable objetivo, permite interpretar de forma directa el error medio cometido por el modelo (Hyndman y Koehler, 2006). En problemas de predicción energética, esta propiedad facilita la lectura operativa del resultado, ya que el error puede relacionarse con unidades físicas de generación.

El R^2 puede formularse como:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.14)$$

donde $\bar{y}$ representa la media de los valores observados. Esta métrica compara la variabilidad no explicada por el modelo con la variabilidad total de la serie respecto a su media. Valores próximos a 1 indican mayor capacidad explicativa, mientras que valores bajos reflejan una menor reducción del error frente a una referencia basada en la media. No obstante, el uso de R^2 en predicción temporal requiere cautela; un valor elevado puede reflejar buen ajuste global de la serie sin garantizar un desempeño adecuado fuera de muestra, especialmente en presencia de tendencia, autocorrelación o cambios de régimen. Por esta razón, R^2 se utilizará como medida complementaria en caso necesario, mientras que la comparación principal entre modelos se apoyará en métricas directas de error, especialmente MAE y RMSE. Esta combinación permite evaluar tanto la magnitud media del error como la presencia de desviaciones puntuales elevadas, dos aspectos relevantes para valorar la utilidad predictiva de los modelos considerados (Hyndman y Koehler, 2006).

3.5. Modelos de referencia para la evaluación del desempeño predictivo

La evaluación de modelos de predicción requiere la definición previa de referencias simples y transparentes. Estas referencias, conocidas como modelos *baseline*, permiten establecer un umbral mínimo de desempeño frente al cual comparar modelos de mayor complejidad. En series temporales, esta comparación adquiere especial relevancia, ya que reglas sencillas basadas en la persistencia o en promedios recientes pueden ofrecer resultados competitivos cuando la serie presenta estabilidad, autocorrelación o baja variabilidad. Por ello, la literatura de predicción recomienda contrastar cualquier propuesta metodológica con modelos simples antes de atribuir valor práctico a arquitecturas más complejas (Hyndman y Koehler, 2006). En este trabajo se utilizará, en primer lugar, un modelo *naïve* como referencia mínima. Este enfoque parte de una regla de predicción simple basada en la información histórica más reciente de la serie. En su forma más habitual, el método *naïve* asume persistencia inmediata y utiliza el último valor observado como predicción del siguiente periodo. Para un horizonte de un paso, esta formulación se expresa como:

$$\hat{y}_{t+1|t} = y_t \quad (3.15)$$

donde y_t representa el último valor observado y $\hat{y}_{t+1|t}$ la predicción realizada para el periodo siguiente con la información disponible hasta t . En el contexto de la generación eléctrica a partir de biogás, esta regla equivale a asumir una continuidad inmediata de la producción entre periodos consecutivos. Su utilidad como referencia procede precisamente de su simplicidad, ya que cualquier modelo más complejo debería mejorar este umbral para justificar su uso. Además de esta formulación basada en el último valor observado, puede considerarse una variante basada en media móvil. En este caso, la predicción se calcula como el promedio de una ventana reciente de observaciones, de tamaño k , según la expresión:

$$\hat{y}_{t+1|t} = \frac{1}{k} \sum_{j=0}^{k-1} y_{t-j} \quad (3.16)$$

Esta estrategia suaviza fluctuaciones puntuales y proporciona una referencia adecuada cuando la serie presenta estabilidad local con pequeñas oscilaciones de corto plazo. La elección del tamaño de la ventana debe responder a la granularidad temporal de los datos y al comportamiento observado en la serie. Una ventana demasiado corta reproduce casi por completo la persistencia inmediata, mientras que una ventana excesivamente larga puede diluir variaciones recientes relevantes para la predicción.

Junto a los modelos de referencia *naïve* y media móvil, se incorporará un modelo estadístico tradicional basado en estructuras autorregresivas. En particular, se considerará el modelo SARIMA. Este enfoque amplía el modelo ARIMA, mediante la incorporación explícita de componentes estacionales. Su formulación permite representar dependencia respecto a valores pasados, diferenciación para tratar tendencias o no estacionariedad, y patrones repetitivos asociados a una periodicidad determinada (Box et al., 2015; Hyndman y Athanasopoulos, 2021).

El modelo SARIMA suele expresarse como $\text{SARIMA}(p, d, q)(P, D, Q)_s$, donde p indica el orden autorregresivo no estacional, d el grado de diferenciación no estacional y q el orden de media móvil no estacional. Los parámetros P , D y Q cumplen una función análoga para la parte estacional, mientras que s representa la periodicidad de la serie. Esta estructura resulta adecuada cuando la variable presenta patrones temporales recurrentes, por ejemplo ciclos semanales, mensuales o estacionales, además de dependencia de corto plazo. A diferencia del modelo *naïve*, SARIMA incorpora una representación explícita de la estructura temporal de la serie bajo un marco lineal. Por tanto, constituye una referencia más exigente, ya que permite evaluar el desempeño alcanzable mediante un modelo estadístico clásico bien especificado. Su inclusión resulta pertinente en este trabajo porque la generación eléctrica a partir de biogás puede presentar persistencia, tendencia suave o patrones periódicos asociados a la operación de las plantas y a la disponibilidad de sustratos.

La incorporación conjunta del modelo *naïve*, la media móvil y SARIMA permite establecer distintos niveles de comparación. El *naïve* fija un umbral mínimo basado en persistencia temporal, la media móvil aporta una referencia suavizada basada en el comportamiento reciente y SARIMA proporciona un estándar estadístico capaz de capturar dependencia temporal y estacionalidad. Frente a estos modelos se evaluarán las arquitecturas RNN, LSTM y GRU. De este modo, la comparación permitirá determinar si la mayor complejidad de los modelos basados en redes neuronales se traduce en una reducción efectiva del error de predicción fuera de muestra.

Capítulo 4

Desarrollo experimental y resultados

En este capítulo se presentan los resultados obtenidos tras la aplicación de la metodología propuesta al problema de predicción de la generación eléctrica a partir de biogás en Galicia. En primer lugar, se describe el flujo metodológico seguido durante la investigación, con el fin de contextualizar las distintas etapas del análisis y facilitar la interpretación de los resultados posteriores. A continuación, se presentan los resultados correspondientes a la adquisición y preprocesamiento de los datos, el análisis descriptivo de la serie temporal y el desarrollo de los modelos predictivos.

Con el objetivo de favorecer la reproducibilidad del estudio, el código utilizado para la adquisición de datos, el preprocesamiento, la implementación de los modelos y la generación de resultados se encuentra disponible en un repositorio público de GitHub, accesible a través del siguiente enlace: https://github.com/candelastb/TFG_BusinessAnalytics_BiogasPrediction. Este repositorio contiene los scripts empleados durante el desarrollo del trabajo y permite replicar los experimentos realizados. Finalmente, se comparan los resultados obtenidos por los modelos de referencia y por las arquitecturas basadas en redes neuronales recurrentes mediante las métricas definidas en el capítulo anterior.

4.1. Diseño metodológico

El desarrollo empírico del trabajo se organiza en un diseño metodológico secuencial, orientado a transformar los registros originales de generación en una comparación homogénea de modelos predictivos. Como se muestra en la Figura 4.1, el proceso comienza con la adquisición de los datos diarios de generación eléctrica a partir de biogás en las cuatro provincias gallegas descargados desde el Sistema de Información del Operador del Sistema (ESIOS). A partir de estos archivos provinciales se construye una serie temporal agregada para Galicia, expresada en MWh diarios, que constituye la variable objetivo del estudio. La segunda etapa comprende la preparación de la base de datos. En esta fase se normalizan las marcas temporales, se revisa la cobertura de cada provincia, se identifican posibles ausencias,

duplicados o inconsistencias y se define el criterio de agregación regional. Este paso permite obtener una serie temporal continua y comparable a lo largo del periodo analizado, evitando que diferencias de formato o cobertura entre archivos condicionen la etapa de modelado.

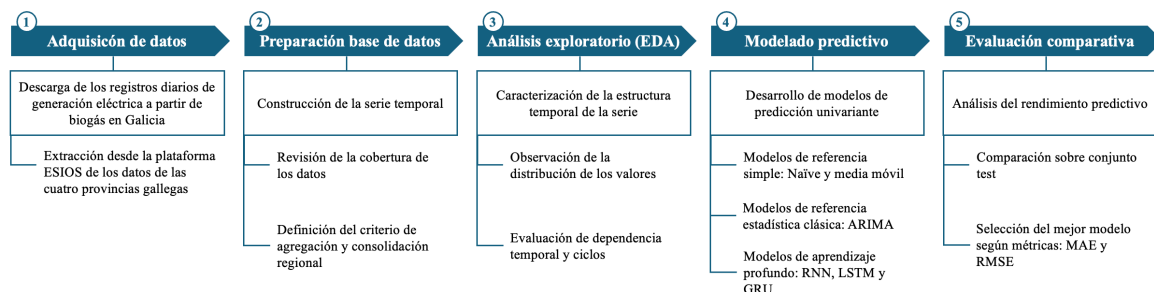


Figura 4.1: Esquema metodológico del trabajo. Fuente: Elaboración propia.

Una vez construida la serie regional, se realiza el análisis exploratorio de los datos. Esta fase examina la evolución temporal de la generación, la distribución de los valores, la presencia de observaciones atípicas y los patrones de variación asociados al calendario. También se estudia la dependencia temporal mediante herramientas de autocorrelación, con el fin de orientar la definición de ventanas de entrada y la elección de modelos con capacidad para representar persistencia, estacionalidad y dependencia entre observaciones. Posteriormente, la fase de modelado predictivo se plantea como una comparación progresiva entre enfoques de distinta complejidad. En primer lugar, se implementan modelos de referencia simples, concretamente el modelo *naïve* y la media móvil, para establecer un umbral mínimo de desempeño. A continuación, se ajusta un modelo SARIMA como referencia estadística clásica basada en dependencia temporal y estacionalidad. Finalmente, se entrenan las arquitecturas recurrentes RNN, LSTM y GRU, aplicando criterios homogéneos de partición temporal, búsqueda de hiperparámetros, entrenamiento y validación. La última etapa corresponde a la evaluación comparativa. La serie se divide cronológicamente en conjuntos de entrenamiento, validación y prueba, respetando el orden temporal de las observaciones para evitar filtración de información futura. El desempeño de los modelos se mide mediante MAE y RMSE, ambas expresadas en MWh, lo que permite interpretar los errores en términos de generación diaria y comparar de forma consistente el aporte predictivo de cada enfoque.

Los datos utilizados en este trabajo proceden de ESIOS, la plataforma de acceso público gestionada por Red Eléctrica de España (REE), que centraliza información operativa del sistema eléctrico nacional. El indicador seleccionado corresponde a la generación medida de biogás con desagregación geográfica a nivel provincial. Esta variable recoge la generación eléctrica diaria producida a partir de biogás y se expresa en megavatios-hora (MWh). Al tratarse de registros procedentes de medidas reportadas al operador del sistema, los valores reflejan la producción efectivamente incorporada al sistema eléctrico durante cada jornada.

Para construir la base de datos del estudio se realizaron cuatro descargas individuales, una por cada provincia gallega: A Coruña, Lugo, Orense y Pontevedra. El periodo de observación

abarca desde el 6 de abril de 2016 hasta el 22 de marzo de 2026, lo que permite disponer de una serie diaria suficientemente extensa para el análisis exploratorio y la posterior estimación de modelos predictivos. Cada archivo fue descargado en formato CSV y conserva la estructura estándar de los registros de ESIOS, compuesta por las variables `id`, `name`, `geoid`, `geoname`, `value` y `datetime`. La Tabla 4.1 resume el contenido de cada campo y su función dentro de la base original.

Tabla 4.1: Variables contenidas en los archivos CSV descargados de ESIOS.

Variable	Descripción	Tipo
<code>id</code>	Identificador numérico interno del indicador en ESIOS.	Numérica entera
<code>name</code>	Nombre descriptivo del indicador descargado.	Cadena de texto
<code>geoid</code>	Código geográfico asociado a la provincia.	Numérica entera
<code>geoname</code>	Nombre de la provincia a la que corresponde el registro.	Cadena de texto
<code>value</code>	Generación eléctrica diaria a partir de biogás, expresada en MWh.	Numérica continua
<code>datetime</code>	Fecha y hora de la observación en formato ISO 8601 con información de zona horaria.	Fecha–tiempo

Antes de construir la base de datos final, se revisaron los cuatro archivos descargados con el fin de verificar su integridad y detectar posibles irregularidades con impacto sobre el análisis posterior. La revisión inicial mostró que ninguno de los archivos contenía valores nulos ni registros duplicados. Por tanto, no fue necesario aplicar procedimientos de imputación ni depuración sobre los valores individuales de generación, lo que permitió conservar los registros originales reportados por ESIOS. Sin embargo, el análisis de la cobertura temporal conjunta mostró diferencias entre provincias. Aunque los cuatro archivos fueron descargados para el mismo intervalo temporal, no todas las fechas presentaban registros simultáneos para las provincias. Para cuantificar esta asimetría, se calculó en cada fecha el número de provincias con dato disponible. La distribución obtenida se recoge en la Tabla 4.2.

Tabla 4.2: Distribución de fechas según el número de provincias con dato registrado.

Provincias con dato	Número de fechas
1	883
2	2012
3	724
4	19

De las 3638 fechas del periodo analizado, únicamente 19 cuentan con registro simultá-

neo en las cuatro provincias. En cambio, 2012 fechas presentan datos en dos provincias y 883 fechas solo en una. Esta distribución evidencia una cobertura provincial heterogénea, compatible con la desigual implantación territorial de las instalaciones de biogás en Galicia. En provincias con menor capacidad instalada o menor actividad registrada, la ausencia de dato puede reflejar jornadas sin generación reportada, producción nula o valores situados por debajo del umbral de comunicación al operador del sistema. Bajo esta interpretación, la falta de registro no se considera necesariamente un error de medición, sino una manifestación de la propia estructura productiva regional. Trabajar de forma independiente con series provinciales de cobertura tan dispar podría introducir discontinuidades y dificultar la comparación temporal entre territorios. Por este motivo, se optó por construir una única serie temporal agregada para Galicia, sumando en cada fecha los valores disponibles en los cuatro archivos provinciales. Esta decisión permite integrar toda la información observada y mantener una frecuencia diaria homogénea durante el periodo de estudio. La serie resultante representa la generación diaria agregada de biogás en Galicia bajo el supuesto de que las ausencias provinciales equivalen a ausencia de generación registrada en esa jornada. No obstante, la imposibilidad de distinguir con total certeza entre generación nula, falta de reporte o ausencia de instalación activa constituye una limitación metodológica que será retomada en la discusión final del trabajo. Una vez definido el criterio de agregación regional, se procedió a la construcción de la base de datos utilizada en el análisis. Este proceso se realizó en Python mediante la librería pandas. En una primera etapa se seleccionaron exclusivamente las variables `datetime` y `value` de cada archivo provincial, dado que el resto de columnas contenían información descriptiva o identificadores administrativos sin utilidad directa para la modelización temporal. Posteriormente, la variable `value` fue renombrada como `generacion_biogas_mwh`, con el propósito de reflejar explícitamente tanto la magnitud representada como su unidad de medida.

La variable temporal requirió un tratamiento adicional debido a la coexistencia de distintos desplazamientos horarios asociados a los cambios estacionales. Los archivos originales almacenan las fechas en formato ISO 8601 con información de zona horaria, utilizando un desplazamiento de +01:00 durante el horario de invierno y +02:00 durante el horario de verano. Esta circunstancia impide una conversión directa a un formato temporal uniforme. Para resolver esta situación, todas las marcas temporales fueron interpretadas inicialmente en tiempo universal coordinado (UTC) y posteriormente convertidas a la zona horaria correspondiente a Madrid. A continuación, se normalizó la componente horaria fijando la medianoche como referencia temporal para cada registro. Este procedimiento garantiza que todas las observaciones correspondientes a una misma jornada compartan exactamente la misma fecha, independientemente de la hora específica almacenada en el archivo original. Conviene señalar que el indicador descargado desde ESIOS ya presenta frecuencia diaria, por lo que esta transformación no implica agregación temporal ni pérdida de información, sino únicamente una normalización del formato de almacenamiento. Finalmente, se eliminó

la información de zona horaria y la variable quedó almacenada como `datetime64`, formato requerido para las operaciones de indexación, agrupación y análisis temporal realizadas posteriormente.

Una vez estandarizadas las variables de los cuatro archivos provinciales, se procedió a la construcción de la serie regional. Para ello, todas las observaciones fueron agrupadas por fecha y se calculó la suma de los valores de generación disponibles en cada jornada. Bajo este criterio, las fechas con información procedente de un subconjunto de provincias incorporan únicamente las contribuciones registradas para dichas provincias. Este procedimiento mantiene el criterio metodológico previamente descrito para las ausencias provinciales y permite obtener una serie diaria continua a escala regional. El resultado de este proceso es una serie temporal univariante que representa la generación eléctrica diaria agregada a partir de biogás en Galicia durante el periodo comprendido entre abril de 2016 y marzo de 2026. La base de datos final contiene 3638 observaciones con frecuencia diaria y dos variables: la fecha de observación y la generación agregada expresada en MWh. Su estructura se presenta en la Tabla 4.3. Esta serie constituye la unidad de análisis sobre la que se desarrollan las etapas posteriores del trabajo, incluyendo el análisis exploratorio, la partición temporal de los datos y el entrenamiento, validación y evaluación de los distintos modelos predictivos considerados

Tabla 4.3: Estructura de la base de datos de generación de biogás en Galicia.

<code>datetime</code>	<code>generacion_biogas_mwh</code>
2016-04-06	78,43
2016-04-07	78,66
2016-04-08	84,38
...	...

4.2. Análisis exploratorio de los datos

El análisis exploratorio tiene como objetivo caracterizar la serie temporal antes de la estimación de los modelos predictivos. Esta fase permite detectar irregularidades en los datos, comprender su distribución y variabilidad, identificar patrones estacionales e interanuales, y obtener indicios sobre la estructura temporal de la serie que orientarán el diseño de los modelos de DL.

4.2.1. Caracterización general de la serie y tratamiento de valores atípicos

La Figura 4.2 presenta la evolución diaria de la generación eléctrica a partir de biogás en Galicia durante el periodo comprendido entre abril de 2016 y marzo de 2026. En términos

generales, la serie muestra un nivel de generación relativamente estable, con oscilaciones moderadas en la mayor parte del horizonte temporal. Sin embargo, se observa un pico de magnitud excepcional en torno al año 2020, cuyo valor se aleja de forma considerable del rango habitual de la serie y altera la escala visual de la representación. Esta anomalía justifica un análisis específico de valores atípicos antes de profundizar en la caracterización temporal de los datos.

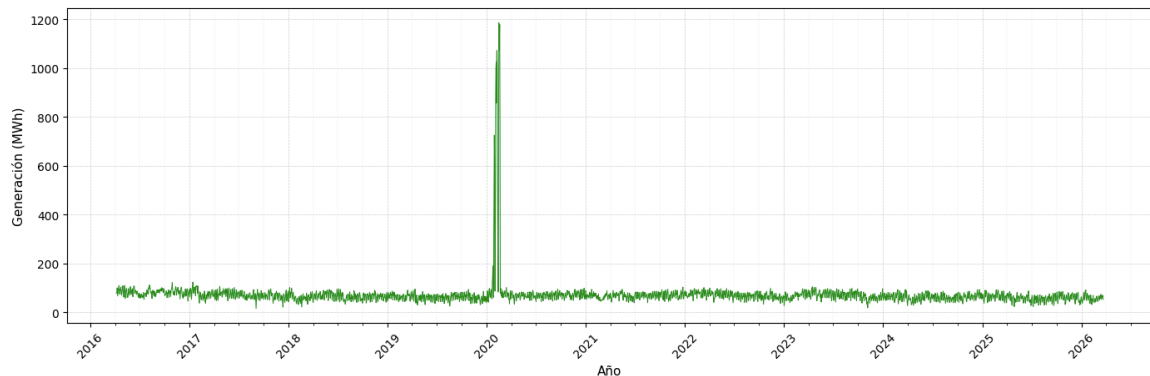


Figura 4.2: Evolución diaria de la generación eléctrica a partir de biogás en Galicia. Fuente: Elaboración propia.

La Tabla 4.4 recoge los principales estadísticos descriptivos de la serie original. La generación media diaria se sitúa en 71,80 MWh, mientras que la mediana alcanza 67,96 MWh. La proximidad entre ambos valores apunta a una distribución concentrada en torno a niveles medios de generación, aunque afectada por valores extremos. Esta interpretación se refuerza al comparar el percentil 75, situado en 78,12 MWh, con el valor máximo de 1185,37 MWh. La distancia entre ambos confirma que el máximo no representa el comportamiento ordinario de la serie, sino una observación excepcional con capacidad para distorsionar medidas sensibles a valores extremos, como la media y la desviación típica. La desviación típica, de 61,57 MWh, debe interpretarse teniendo en cuenta esta anomalía, ya que incorpora la variabilidad inducida por el pico detectado.

Tabla 4.4: Estadísticos descriptivos de la serie de generación de biogás, MWh/día.

Estadístico	Valor (MWh)
Observaciones	3638
Media	71,80
Desviación típica	61,57
Mínimo	15,62
Percentil 25	57,29
Mediana	67,96
Percentil 75	78,12
Máximo	1185,37

La comparación entre los cuantiles y el valor máximo permite concluir que la serie requiere un tratamiento diferenciado según la finalidad del análisis. Para la descripción exploratoria, la presencia del pico puede ocultar la variabilidad habitual de la generación y dificultar la interpretación de los patrones temporales. Con el fin de evaluar de forma objetiva la magnitud de esta anomalía, se aplicaron dos procedimientos ampliamente utilizados para la detección de valores atípicos: el criterio basado en el Z-score y el método del rango intercuartílico (IQR). El Z-score cuantifica la distancia de cada observación respecto a la media en unidades de desviación típica. Adoptando el umbral convencional de $|z| > 3$, se identificaron 20 observaciones atípicas, todas ellas concentradas en el intervalo temporal asociado al pico previamente observado. De forma complementaria, el método IQR define como atípicos los valores situados fuera del intervalo comprendido entre $Q_1 - 1,5 \cdot IQR$ y $Q_3 + 1,5 \cdot IQR$. Bajo este criterio se detectaron 19 observaciones, incluyendo algunos valores elevados registrados en los últimos años de la serie que, aunque extremos desde un punto de vista estadístico, parecen compatibles con la variabilidad propia del proceso de generación. La Figura 4.3 muestra la localización de los valores identificados mediante ambos procedimientos sobre la serie original.

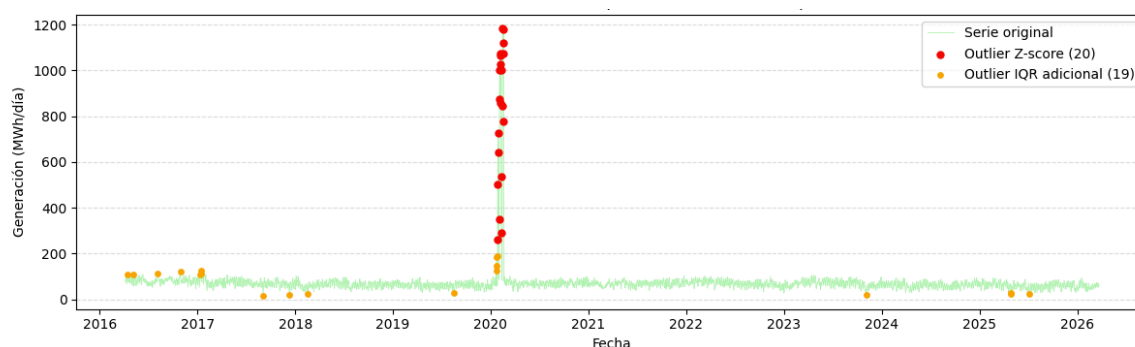


Figura 4.3: Identificación de valores atípicos sobre la serie temporal. Fuente: Elaboración propia.

La concentración espacial y temporal de estas observaciones sugiere que la anomalía no responde al comportamiento ordinario de la generación regional. Una revisión de los archivos originales permitió localizar el origen del pico en registros concretos asociados a la provincia de A Coruña. Dada su magnitud y su aparición en un intervalo temporal muy reducido, resulta razonable considerar que estos valores podrían estar relacionados con incidencias en el proceso de registro o reporte de la información. No obstante, al no disponerse de documentación oficial que confirme esta hipótesis, dicha interpretación debe entenderse como una explicación plausible y no como una evidencia concluyente. Esta incertidumbre condiciona la estrategia de tratamiento adoptada. En el análisis exploratorio se optó por corregir las observaciones identificadas mediante el criterio Z-score, dado que su presencia altera de forma significativa la escala de las representaciones gráficas y dificulta la caracterización de la dinámica habitual de la serie. Por el contrario, en la fase de modelado predictivo se

mantuvo la serie original sin modificaciones. Esta decisión responde al interés de evaluar el comportamiento de los modelos ante datos observados en condiciones reales, incluyendo posibles anomalías o perturbaciones presentes en los registros históricos. De este modo, la comparación entre modelos no solo considera su capacidad predictiva, sino también su sensibilidad frente a observaciones extremas. La corrección aplicada durante el análisis exploratorio consistió en sustituir las observaciones identificadas como atípicas por valores ausentes (NaN) y reconstruir posteriormente dichos puntos mediante interpolación lineal. Este procedimiento preserva la continuidad temporal de la serie y permite suavizar alteraciones puntuales sin introducir transformaciones adicionales sobre su estructura global (Hyndman y Athanasopoulos, 2021). A partir de este punto, los análisis descriptivos se realizan sobre la serie corregida, mientras que las etapas de entrenamiento y evaluación de los modelos utilizan la serie original.

4.2.2. Distribución y patrones temporales de generación

El histograma de frecuencias y la curva *Kernel Density Estimation*, de la Figura 4.4 permiten evaluar la distribución de los valores diarios de generación sobre la serie corregida. El histograma muestra la frecuencia de observaciones en distintos intervalos de generación, mientras que la curva ofrece una representación suavizada de la densidad. La lectura conjunta de ambas representaciones permite evaluar la concentración de valores, la simetría de la distribución y la posible presencia de más de un régimen de operación.

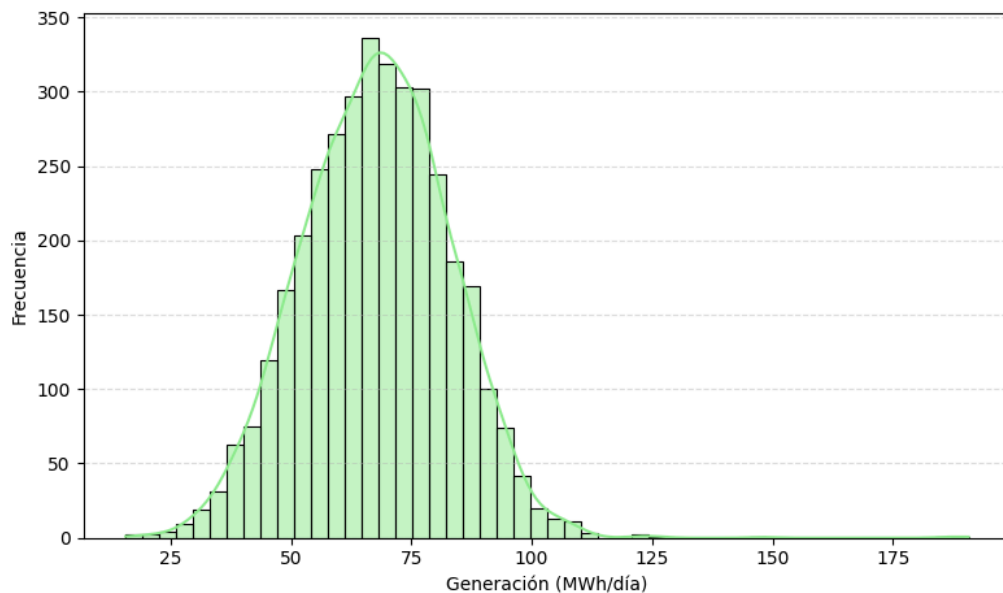


Figura 4.4: Histograma y estimación de densidad por núcleo de la generación diaria de biogás en Galicia. Fuente: Elaboración propia.

La Figura 4.4 muestra una distribución unimodal, con la mayor concentración de observaciones en torno a los 65–70 MWh diarios. La asimetría positiva moderada, estimada en 0,27,

indica una cola derecha ligeramente más extensa, asociada a días de generación elevada menos frecuentes y más dispersos que los valores situados en torno al centro de la distribución. La curtosis, con un valor de 1,95, se sitúa por debajo del valor de referencia de una distribución normal, lo que apunta a una distribución platicúrtica con colas relativamente ligeras tras el tratamiento del pico de 2020. Esta distribución refleja una operativa relativamente estable, concentrada alrededor de un rango habitual de generación y sin indicios de comportamientos claramente diferenciados. A partir de esta caracterización de la distribución global, se analizó si la generación presenta patrones regulares asociados al calendario. Para ello, se elaboraron diagramas de caja sobre la serie corregida, agrupando las observaciones por mes y por día de la semana. Cada diagrama resume la mediana, el IQR y los valores extremos de la generación diaria dentro de cada grupo temporal, lo que permite comparar no solo el nivel central de producción, sino también su dispersión. La Figura 4.5 muestra que las diferencias entre meses son reducidas. Las medianas se mantienen en un intervalo estrecho, cercano a los 65–70 MWh diarios, y los IQR presentan una amplitud similar en la mayor parte del año. No se observa una separación sistemática entre meses de invierno y verano, por lo que la generación de biogás en Galicia no parece estar condicionada de forma marcada por el ciclo climático anual. Esta estabilidad mensual resulta coherente con la naturaleza del recurso, dado que la digestión anaerobia depende en mayor medida de la disponibilidad de materia orgánica, la capacidad instalada y las condiciones de operación de las plantas que de variaciones estacionales directas.

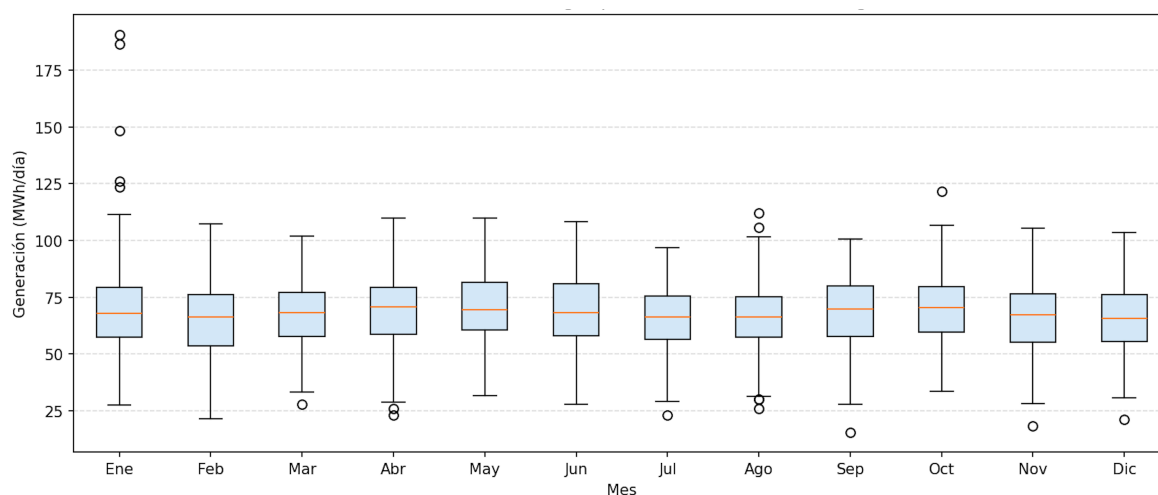


Figura 4.5: Distribución de la generación diaria de electricidad a partir de biogás por mes del año. Fuente: Elaboración propia.

La agregación por día de la semana muestra un patrón distinto. Como se observa en la Figura 4.6, los lunes y domingos presentan medianas inferiores, situadas aproximadamente entre 50 y 55 MWh, mientras que jueves, viernes y sábado registran los niveles más altos, en torno a 75–80 MWh. Esta diferencia sugiere la presencia de un componente semanal en la serie, posiblemente asociado a prácticas operativas, criterios de despacho, ritmos de recogida

de residuos o dinámicas de alimentación de los digestores vinculadas al calendario laboral. A diferencia del patrón mensual, la variación semanal sí introduce una estructura temporal que deberá ser considerada en la fase de modelado predictivo.

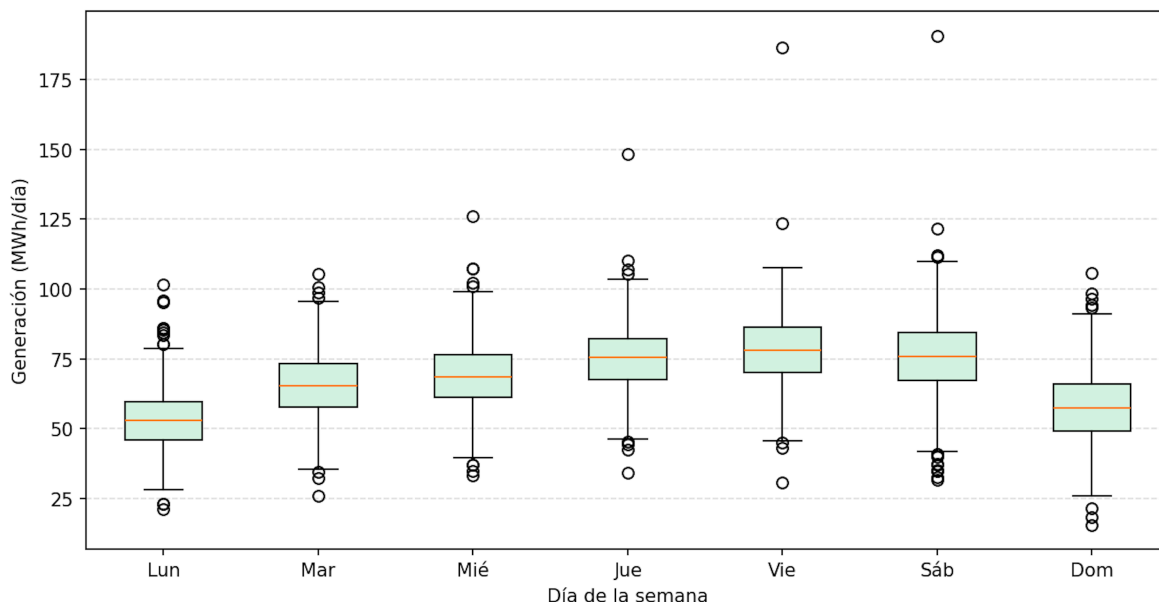


Figura 4.6: Distribución de la generación diaria de electricidad a partir de biogás por día de la semana. Fuente: Elaboración propia.

Para completar el análisis de patrones temporales, se evaluó la evolución interanual de la generación. Este análisis se realizó únicamente sobre los años con doce meses completos de información, con el fin de evitar sesgos derivados de los periodos incompletos al inicio y al final de la serie. Para ello, se utilizaron dos representaciones complementarias: la distribución de la generación mensual total por año y la media diaria anual. La Figura 4.7 muestra la distribución de la generación mensual total para el periodo 2017–2025. La evolución interanual no sigue una trayectoria lineal uniforme, sino una secuencia de fases diferenciadas. En primer lugar, se observa una reducción desde los niveles más altos de 2017, con una mediana próxima a 2200 MWh mensuales, hasta un mínimo en 2019, con una mediana aproximada de 1850 MWh. Posteriormente, la serie presenta una recuperación entre 2020 y 2023, con medianas que retornan al entorno de 2100–2200 MWh mensuales y alcanzan su nivel más elevado en 2023. En el tramo final, correspondiente a 2024 y 2025, se aprecia una nueva caída, con valores mensuales similares o inferiores a los registrados en 2019. La persistencia de esta reducción requeriría observaciones posteriores al periodo analizado, por lo que su interpretación debe mantenerse como una tendencia reciente dentro de la ventana disponible.

La Figura 4.8 complementa la lectura anterior mediante la media diaria por año, lo que permite comparar de forma directa el nivel medio de generación entre ejercicios. Esta representación reproduce la misma estructura temporal observada en el boxplot anual: descenso inicial, recuperación intermedia y caída en los años más recientes. Dado que la base dispo-

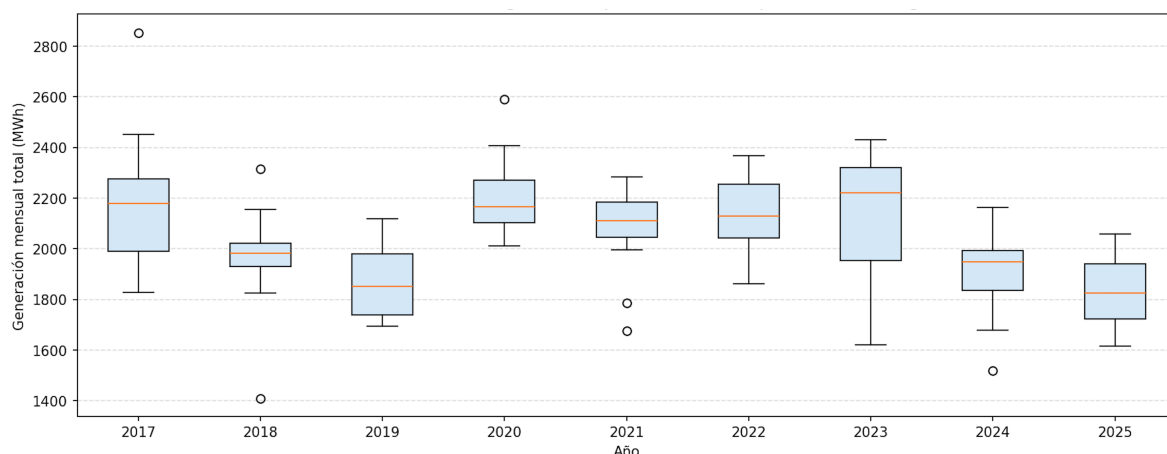


Figura 4.7: Variación mensual de la generación por año completo (2017 - 2025). Fuente: Elaboración propia.

nible contiene únicamente generación agregada, no es posible atribuir estas variaciones a factores específicos. Entre las posibles explicaciones se encuentran cambios en la disponibilidad o composición de los sustratos, modificaciones operativas en las plantas, variaciones en el despacho de energía o envejecimiento de las instalaciones. Estas interpretaciones se plantean como hipótesis compatibles con la evolución observada, pero no pueden verificarse directamente con las variables disponibles en este trabajo.

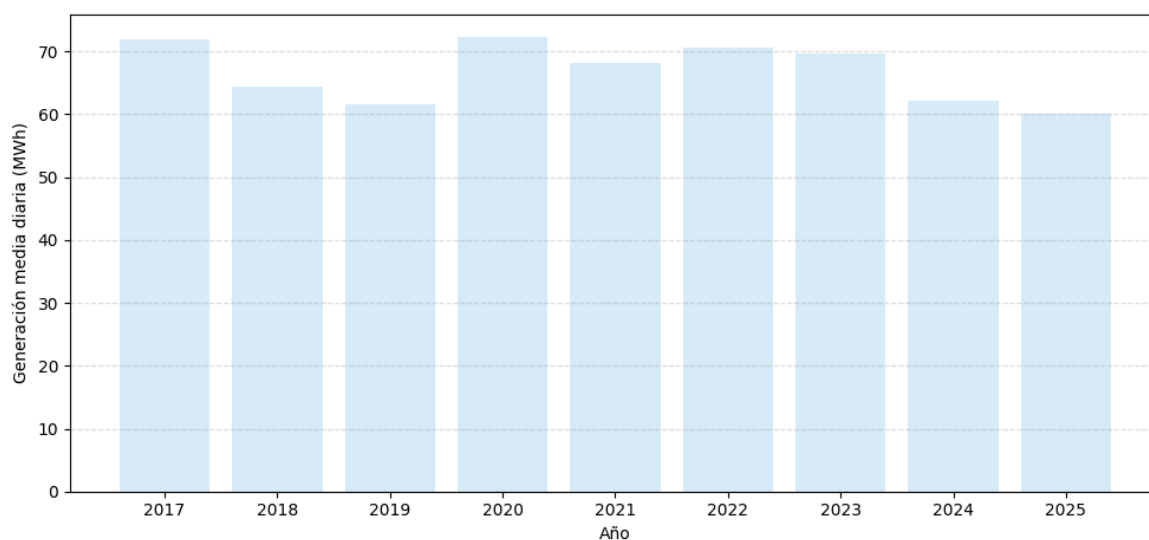


Figura 4.8: Generación media diaria por año completo (2017 - 2025). Fuente: Elaboración propia.

4.2.3. Tendencia, estacionalidad y dependencia temporal

Con el fin de analizar la estructura temporal de la serie más allá de los patrones descriptivos observados anteriormente, se analizaron sus componentes de tendencia y estacionalidad

mediante dos herramientas complementarias: la media móvil de 30 días y la descomposición estacional aditiva. Ambas permiten aislar comportamientos de distinta escala temporal y facilitan la identificación de regularidades que pueden resultar relevantes para la modelización predictiva. La media móvil de 30 días, representada en la Figura 4.9, suaviza las fluctuaciones diarias y permite visualizar la evolución de largo plazo de la generación. La trayectoria obtenida muestra una reducción progresiva a lo largo del periodo analizado. Los valores iniciales, situados en torno a 85–90 MWh diarios durante 2016, descienden hasta aproximadamente 65 MWh en 2018 y 2019. Posteriormente se observa una fase de relativa estabilidad entre 2020 y 2023, seguida de una nueva reducción en los años finales de la serie, con niveles próximos a 55–60 MWh diarios. El incremento puntual visible alrededor de 2020 corresponde a observaciones asociadas a la anomalía previamente identificada, aunque su efecto aparece atenuado al considerar una ventana móvil de 30 días.

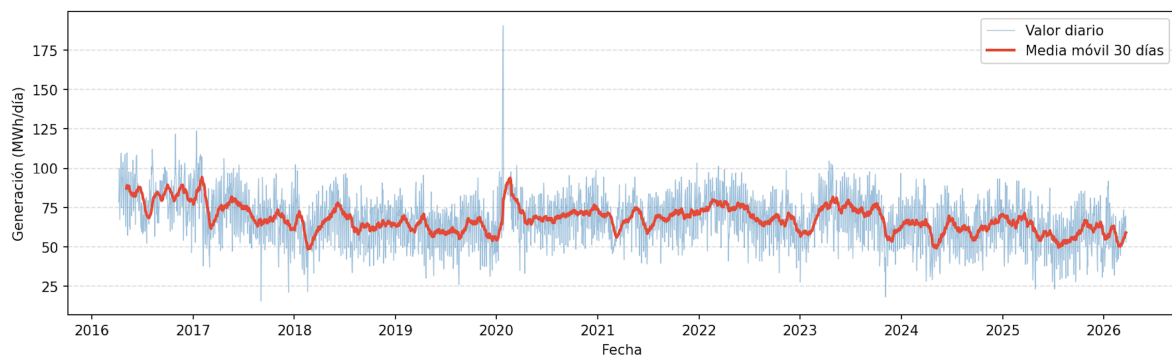


Figura 4.9: Generación diaria y media móvil de 30 días. Fuente: Elaboración propia.

La descomposición estacional aditiva permite evaluar de forma separada las componentes de tendencia, estacionalidad y residuo. Como muestra la Figura 4.10, la componente de tendencia reproduce la evolución descendente observada mediante la media móvil, con una reducción gradual desde valores cercanos a 80 MWh al inicio del periodo hasta aproximadamente 60 MWh en los años finales, interrumpida por una recuperación moderada entre 2020 y 2023. La componente estacional presenta oscilaciones regulares de aproximadamente ± 12 MWh alrededor de la tendencia. Su repetición sistemática a lo largo de toda la serie resulta consistente con el patrón semanal identificado previamente mediante los diagramas de caja por día de la semana. La magnitud de estas oscilaciones es apreciable, aunque inferior a la variabilidad asociada a los cambios de tendencia observados a largo plazo. Por su parte, la componente residual permanece centrada en torno a cero durante la mayor parte del periodo analizado. La principal excepción corresponde al conjunto de residuos positivos concentrados alrededor de 2020, asociados a los registros anómalos identificados anteriormente. Su persistencia en la descomposición confirma que, aun después del tratamiento aplicado, parte del efecto de esta incidencia continúa siendo visible en la estructura residual de la serie.

A partir de la descomposición anterior, se analizó la dependencia temporal de la serie

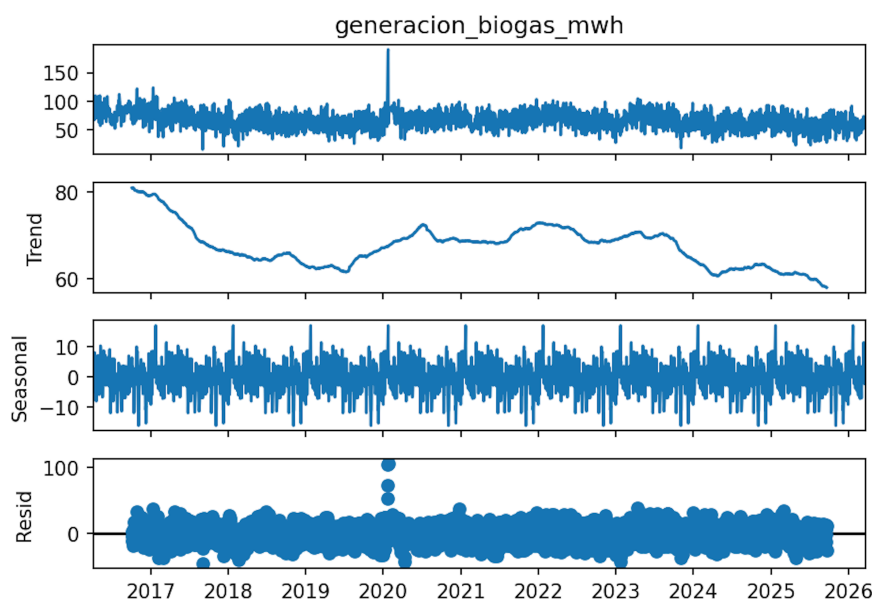


Figura 4.10: Descomposición estacional aditiva de la serie de generación. Fuente: Elaboración propia.

mediante la *Autocorrelation Function* (ACF), y la *Partial Autocorrelation Function* (PACF). Ambas herramientas permiten cuantificar la relación entre la generación observada en una fecha y sus valores pasados en distintos retardos temporales. La ACF mide la correlación total entre la serie y sus retardos, mientras que la PACF estima la correlación directa de cada retardo una vez descontado el efecto de los retardos intermedios. Por tanto, su análisis conjunto representado en la Figura 4.11 resulta útil para identificar dependencias inmediatas, patrones periódicos y posibles ventanas temporales de entrada para los modelos predictivos.

La ACF representada en la Figura 4.11 muestra una autocorrelación elevada en el retardo 1, próxima a 0,70, lo que indica una fuerte persistencia entre días consecutivos. Además, en lugar de presentar una caída monótona, aparecen picos destacados en los retardos múltiplos de siete (7, 14, 21, 28, 35, 42, 49 y 56), en su mayoría por encima de la banda de confianza. Esta pauta confirma la presencia de una periodicidad semanal, dado que la generación de un día guarda relación con la observada en el mismo día de semanas anteriores. La PACF permite precisar esta interpretación al aislar la contribución directa de cada retardo. Los retardos con correlación parcial más claramente significativa son el 1, con un valor aproximado de 0,70, y el 2, con un valor cercano a $-0,29$. En los primeros retardos se aprecian, además, coeficientes positivos de menor magnitud junto a un valor negativo significativo en torno al retardo 8 (aproximadamente $-0,29$). A partir de ese punto, la mayoría de los coeficientes se sitúan dentro o cerca de la banda de confianza. Este patrón indica que la información temporal con mayor contenido predictivo se concentra en los días inmediatamente anteriores, mientras que los picos sostenidos de la ACF en los múltiplos de siete reflejan, en buena medida, la propagación de estas dependencias primarias a través de la estructura semanal de

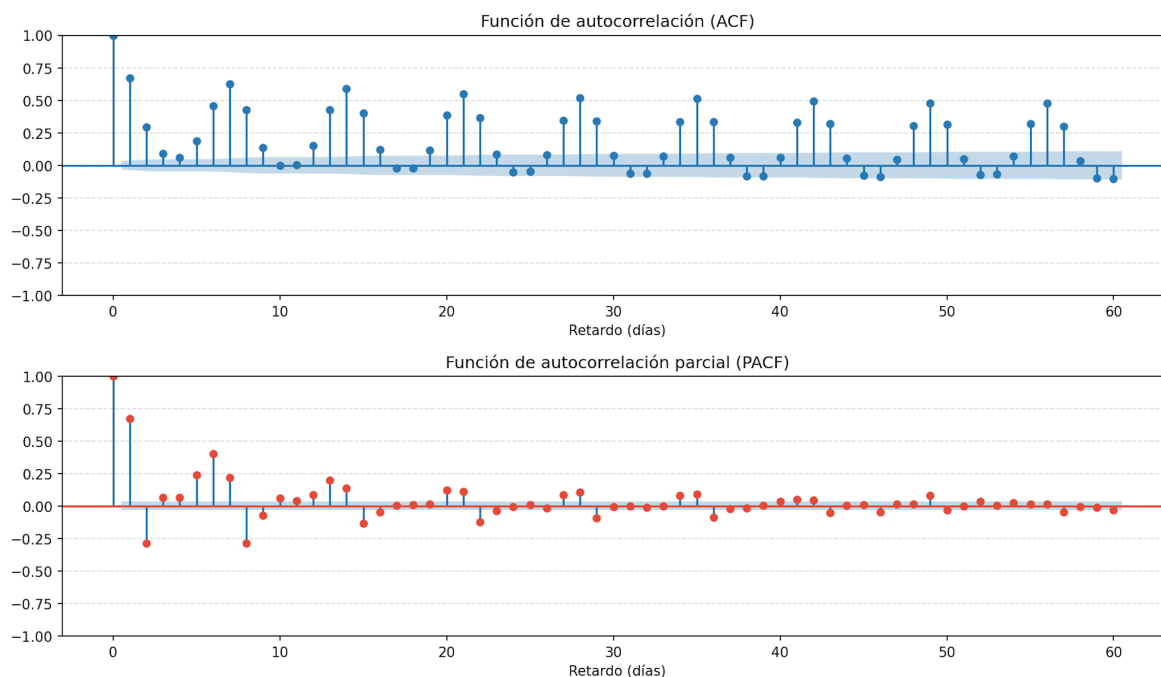


Figura 4.11: Función de autocorrelación y función de autocorrelación parcial de la serie.
Fuente: Elaboración propia.

la serie. Estos resultados tienen implicaciones directas para el diseño del modelado predictivo. La persistencia diaria y la periodicidad semanal justifican el uso de modelos capaces de incorporar información secuencial, como RNN, LSTM y GRU. Además, la concentración de dependencia directa en los primeros retardos y la presencia del ciclo semanal sugieren que ventanas de entrada de 7 a 14 días pueden capturar una parte sustancial de la estructura temporal de la serie, sin incrementar innecesariamente la dimensionalidad del problema.

Como cierre del análisis exploratorio, se incorpora el mapa de calor de la Figura 4.12, que representa la generación mensual total para cada combinación de año y mes en los años con información completa. Esta visualización permite contrastar, en una misma matriz, la existencia de patrones mensuales recurrentes y la evolución interanual de la generación. La lectura por columnas no muestra una estacionalidad mensual marcada. Ningún mes concentra de forma sistemática los valores más altos o más bajos a lo largo de todos los años analizados. Febrero aparece con frecuencia entre los meses de menor generación, con 1828 MWh en 2017, 1410 MWh en 2018, 1704 MWh en 2019 y 1676 MWh en 2021, pero este comportamiento no se reproduce de manera uniforme en todos los ejercicios. Por tanto, la variación mensual presenta un comportamiento irregular y no parece responder a un ciclo climático anual estable. La lectura por filas permite observar con mayor claridad la dinámica interanual. Los años 2020, 2021 y 2022 concentran valores generalmente superiores a 2000 MWh en buena parte de los meses, mientras que 2018, 2019, 2024 y 2025 presentan varios meses por debajo de 1900 MWh. El año 2025 muestra una reducción más homogénea a lo largo de la matriz, consistente con la tendencia descendente identificada previamente mediante la

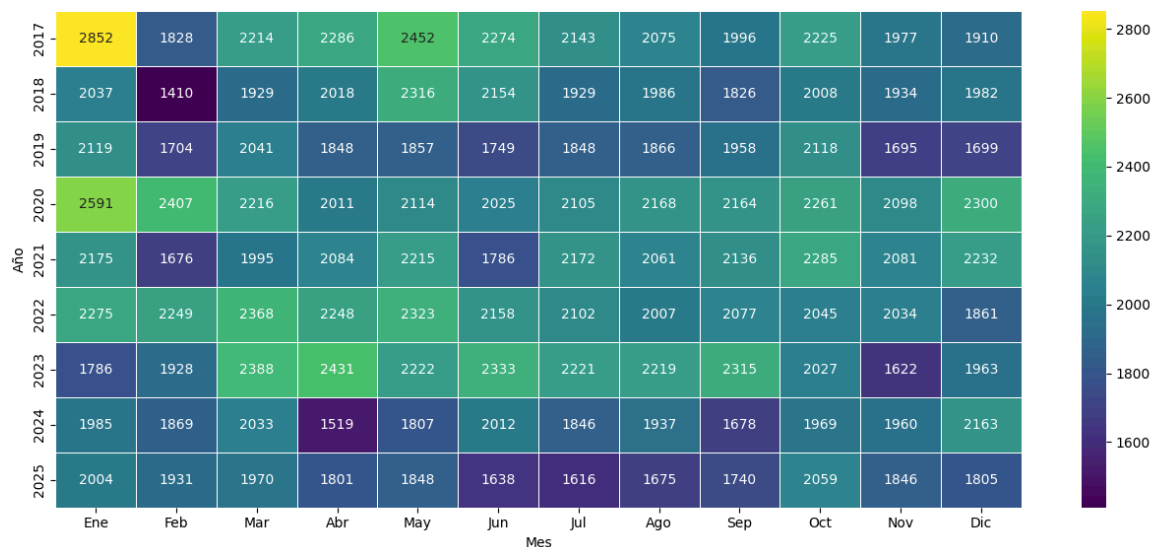


Figura 4.12: Generación mensual total en Galicia, MWh, en años completos. Fuente: Elaboración propia.

media móvil y el análisis interanual. De esta manera, el mapa de calor refuerza la interpretación de que la variabilidad de la serie se explica en mayor medida por cambios entre años que por un patrón mensual estable.

4.3. Modelado Predictivo

Una vez caracterizada la serie temporal y definidas las decisiones de preparación de los datos, se procede a la evaluación de los modelos predictivos orientados a estimar la generación eléctrica diaria a partir de biogás en Galicia. El análisis compara enfoques de distinta naturaleza y complejidad, desde modelos de referencia simples hasta arquitecturas recurrentes de aprendizaje profundo. En primer lugar, se consideran el modelo *naïve* y la media móvil como referencias iniciales de desempeño. A continuación, se incorpora un modelo SARIMA como aproximación estadística clásica capaz de representar dependencia temporal y estacionalidad. Finalmente, se evalúan tres arquitecturas recurrentes, RNN, LSTM y GRU, con el fin de analizar si el aprendizaje de relaciones secuenciales no lineales permite mejorar la precisión predictiva. Todos los modelos se entrenan y evalúan sobre la misma serie temporal, compuesta por 3638 observaciones diarias comprendidas entre el 6 de abril de 2016 y el 22 de marzo de 2026. La partición de los datos se realiza de forma cronológica, asignando el 70 % de las observaciones al conjunto de entrenamiento (2546 registros), el 15 % al conjunto de validación (546 registros) y el 15 % restante al conjunto de prueba (546 registros). Esta división preserva el orden temporal de la serie y evita la utilización de información futura durante el entrenamiento y la selección de modelos (Hyndman y Athanasopoulos, 2021).

Antes del entrenamiento de las arquitecturas recurrentes, la serie se escala mediante *Ro-*

bustScaler. Esta transformación se basa en la mediana y el IQR, por lo que resulta menos sensible a valores atípicos que los procedimientos basados en la media y la desviación típica (Pedregosa et al., 2011). Para evitar filtración de información, los parámetros del escalado se estiman exclusivamente sobre el conjunto de entrenamiento y después se aplican a los conjuntos de validación y prueba. Este preprocesamiento se utiliza únicamente en los modelos neuronales, mientras que los modelos *naïve*, media móvil y SARIMA se evalúan directamente sobre la serie en su escala original. El desempeño predictivo se evalúa mediante el MAE, y el RMSE, ambos expresados en MWh. El MAE mide la desviación media absoluta entre valores observados y predichos, mientras que el RMSE penaliza con mayor intensidad los errores de gran magnitud. La combinación de ambas métricas permite comparar la precisión promedio de los modelos y su sensibilidad ante desviaciones pronunciadas (Hyndman y Koehler, 2006). Las secciones siguientes presentan el proceso de ajuste y evaluación de cada modelo, seguido de una comparación conjunta de los resultados obtenidos con el fin de identificar la alternativa que ofrece el mejor equilibrio entre precisión predictiva y complejidad metodológica.

4.3.1. Modelos de referencia: *naïve* y media móvil

En esta primera fase del modelado se evalúan los dos modelos de referencia definidos en el marco metodológico: el modelo *naïve* y la media móvil. En lugar de aplicar un proceso de entrenamiento, ambos modelos generan sus predicciones directamente a partir de valores históricos recientes de la serie. Por este motivo, su función en esta etapa consiste en fijar un punto de comparación inicial para valorar posteriormente el desempeño de SARIMA y de las arquitecturas recurrentes. El modelo *naïve* se implementó como una regla de persistencia a un paso, de modo que la predicción de cada día corresponde al valor observado en el día inmediatamente anterior. La media móvil, por su parte, se calculó con una ventana de siete días. Esta elección permite incorporar una semana completa de observaciones, en coherencia con la frecuencia diaria de la serie y con el patrón semanal identificado en el análisis exploratorio. A diferencia del modelo *naïve*, que traslada directamente la última observación disponible, la media móvil promedia los valores recientes y genera una predicción más suavizada. Esta característica puede reducir la influencia de fluctuaciones puntuales, aunque también puede introducir cierto retardo ante cambios reales en la generación. Para mantener la comparabilidad con los modelos posteriores, ambos enfoques se evaluaron sobre las mismas particiones cronológicas de entrenamiento, validación y prueba. Aunque estos modelos no tienen parámetros entrenables ni requieren búsqueda de hiperparámetros, se calcularon sus métricas en los tres conjuntos con el fin de establecer una referencia homogénea para el resto del análisis. La Tabla 4.5 recoge los valores de MAE y RMSE obtenidos por ambos modelos.

Tabla 4.5: Métricas de evaluación de los modelos de referencia. *naïve* y media móvil

Conjunto	Naïve		Media móvil (7 días)	
	MAE (MWh)	RMSE (MWh)	MAE (MWh)	RMSE (MWh)
Entrenamiento	12,112	34,127	15,451	54,084
Validación	10,206	12,763	11,371	13,753
Prueba	9,718	12,289	11,021	13,387

El modelo *naïve* alcanza un MAE de 9,718 MWh en el conjunto de prueba. Este valor representa el error medio asociado a una regla de persistencia diaria, es decir, a asumir que la generación del día siguiente será igual a la observada en el día inmediatamente anterior. Por tanto, este resultado fija una referencia inicial exigente para el resto de modelos, ya que cualquier aproximación más compleja deberá reducir este error para justificar su incorporación al proceso predictivo. La comparación entre MAE y RMSE muestra diferencias relevantes en el conjunto de entrenamiento. En el modelo *naïve*, el MAE de entrenamiento es de 12,112 MWh, mientras que el RMSE asciende a 34,127 MWh. En la media móvil, esta diferencia es todavía mayor, con un MAE de 15,451 MWh y un RMSE de 54,084 MWh. Esta separación entre ambas métricas indica la presencia de errores puntuales de gran magnitud en el tramo histórico de entrenamiento, coherentes con los valores atípicos identificados en el análisis exploratorio. Dado que el RMSE eleva los errores al cuadrado antes de promediarlos, su valor aumenta de forma considerable ante desviaciones extremas, mientras que el MAE ofrece una medida menos sensible a este tipo de episodios.

En validación y prueba, los valores de RMSE se reducen de forma marcada para ambos modelos. Este comportamiento sugiere que los tramos más recientes de la serie presentan menor presencia de observaciones extremas y una dinámica diaria más regular que la observada en parte del conjunto de entrenamiento. El mejor desempeño de los modelos de referencia en validación y prueba no debe interpretarse como sobreajuste, ya que ninguno de ellos contiene parámetros entrenables. Más bien refleja diferencias estructurales entre los periodos de la serie utilizados en cada partición. La media móvil de siete días no mejora al modelo *naïve* en ninguno de los conjuntos evaluados. En el conjunto de prueba, su MAE asciende a 11,021 MWh frente a 9,718 MWh del *naïve*, y su RMSE de 13,387 MWh frente a 12,289 MWh en *naïve*. Este resultado indica que, en esta serie, la información contenida en el último valor observado tiene mayor capacidad predictiva que el promedio de la semana anterior. La media móvil reduce la variabilidad de la señal, pero introduce un retraso en la respuesta ante cambios diarios de generación. La inspección gráfica de las predicciones permite complementar la interpretación de las métricas anteriores. La Figura 4.13 representa una porción del conjunto de prueba y compara la serie real con las predicciones generadas por el modelo *naïve* y por la media móvil de siete días. La predicción *naïve* reproduce la trayectoria de la serie

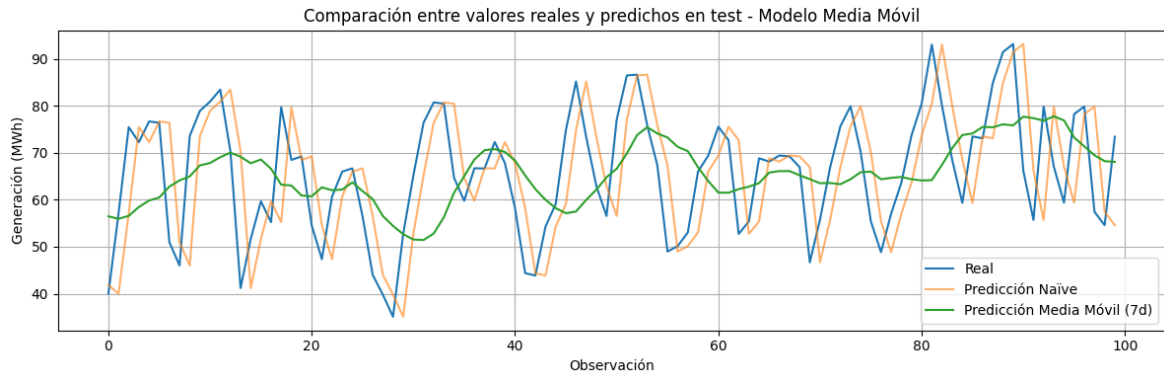


Figura 4.13: Comparación entre los valores reales y las predicciones de los modelos *naïve* y media móvil en el conjunto de prueba. Fuente: Elaboración propia.

con un desplazamiento temporal de un día, por lo que sigue de cerca los cambios de nivel, pero reacciona con retraso ante picos y caídas abruptas. Esta característica explica su buen desempeño relativo en una serie con elevada persistencia diaria, aunque también muestra su limitación para anticipar variaciones repentinas. La media móvil presenta un comportamiento distinto. Su curva es más estable y reduce la amplitud de los picos y valles observados, pero esta suavización implica una pérdida de información sobre la variación diaria de la generación. Por lo tanto, el modelo responde con mayor lentitud a los cambios y tiende a aproximar las predicciones al nivel medio de la ventana semanal. Esta dinámica visual coincide con los resultados de la Tabla 4.5, donde la media móvil obtiene errores superiores al modelo *naïve* en el conjunto de prueba.

Las Figuras 4.14 y 4.15 muestran la distribución del error absoluto en los conjuntos de entrenamiento, validación y prueba para ambos modelos. En los dos casos se observa una dispersión considerablemente mayor en entrenamiento, asociada a la presencia de observaciones extremas identificadas previamente en el análisis exploratorio. Esta diferencia explica la brecha observada entre MAE y RMSE en dicho conjunto y confirma que los episodios más anómalos de la serie se concentran en el tramo histórico utilizado para entrenamiento. De acuerdo a los resultados obtenidos, el modelo *naïve* se adopta como referencia principal para la comparación con los modelos posteriores. Su menor error en validación y prueba, unido a su extrema simplicidad, establece un umbral de desempeño particularmente exigente. De esta manera, el modelo SARIMA, y las arquitecturas RNN, LSTM y GRU deberán superar este nivel de precisión para justificar la complejidad adicional que incorporan.

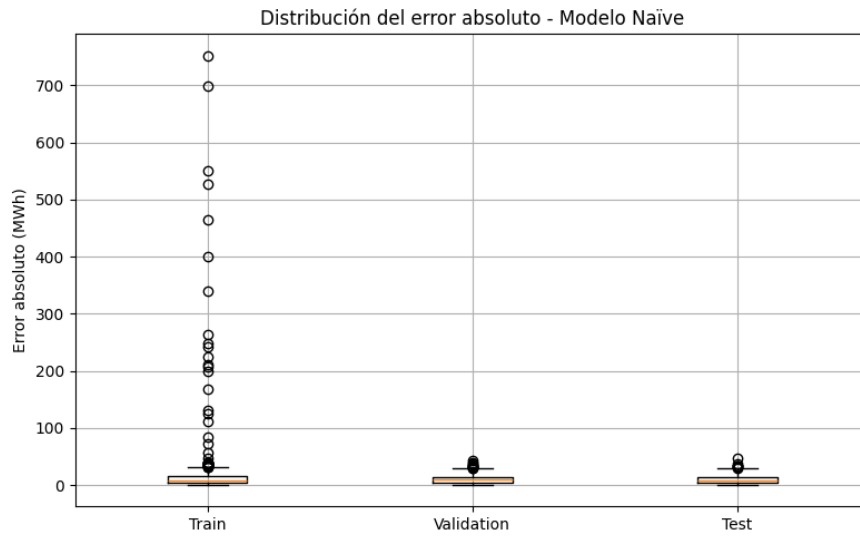


Figura 4.14: Distribución del error absoluto del modelo *naïve* en entrenamiento, validación y prueba. Fuente: Elaboración propia.

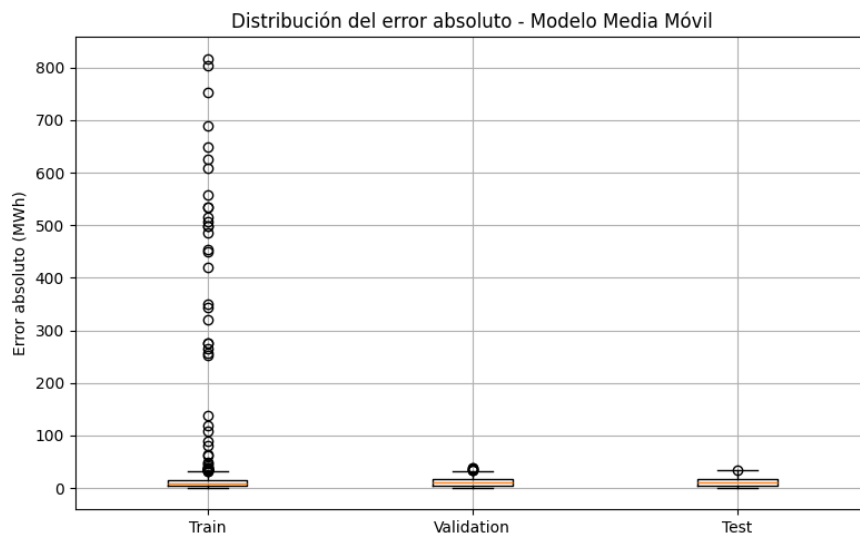


Figura 4.15: Distribución del error absoluto del modelo de media móvil en entrenamiento, validación y prueba. Fuente: Elaboración propia.

4.3.2. Modelo SARIMA

El modelo SARIMA, se incorpora como referencia estadística para la predicción de la serie. Su inclusión permite evaluar el desempeño de un modelo paramétrico lineal con estructura temporal explícita antes de analizar las arquitecturas recurrentes de aprendizaje profundo. La elección de SARIMA, en lugar de un modelo ARIMA no estacional, responde al patrón semanal identificado en el análisis exploratorio. En particular, la función de autocorrelación mostró picos en retardos múltiplos de siete, compatibles con una dependencia periódica

semanal. Por esta razón, se incorporó una componente estacional con periodo $s = 7$, manteniendo la especificación general $\text{SARIMA}(p, d, q)(P, D, Q)_7$. Antes de definir la grilla de parámetros, se evaluó la estacionariedad de la serie original mediante el test de Dickey–Fuller aumentado. Este contraste permite analizar la presencia de raíz unitaria, bajo una hipótesis nula de no estacionariedad frente a una hipótesis alternativa de estacionariedad. El rechazo de la hipótesis nula se produce si el estadístico ADF es más negativo que los valores críticos correspondientes o si el p -valor se sitúa por debajo del nivel de significación considerado. Los resultados del contraste respaldan la estacionariedad de la serie. El estadístico ADF obtenido fue de $-7,693$, valor considerablemente inferior a los valores críticos del 1 %, 5 % y 10 %, situados en $-3,432$, $-2,862$ y $-2,567$, respectivamente. El p -valor asociado fue de $1,40 \times 10^{-11}$, también muy inferior al umbral habitual de 0,05. Ambos criterios conducen al rechazo de la hipótesis nula de raíz unitaria y permiten tratar la serie como estacionaria en niveles. A partir de este resultado, la grilla de búsqueda se acotó a $d \in \{0, 1\}$, conservando $d = 1$ únicamente como alternativa de control ante posibles componentes de tendencia no recogidas por el contraste.

A partir de estos resultados y de la información obtenida en el análisis de autocorrelación, se definió una grilla de búsqueda acotada para el modelo SARIMA. En la parte no estacional se consideraron los valores $p \in \{0, 1, 2\}$, $d \in \{0, 1\}$ y $q \in \{0, 1, 2\}$. La restricción de los órdenes autorregresivo y de media móvil a un máximo de dos retardos responde a la estructura observada en la PACF, donde la correlación parcial directa se concentraba principalmente en los retardos 1 y 2. El resto de retardos, salvo los asociados al ciclo semanal, se situaban dentro o cerca de la banda de confianza. Por tanto, ampliar los órdenes p y q habría incrementado la complejidad del modelo sin una justificación empírica clara. El parámetro d se restringió a los valores $\{0, 1\}$, en coherencia con el resultado del test de Dickey–Fuller aumentado, que indicaba estacionariedad en niveles. Para la componente estacional se consideraron los valores $P \in \{0, 1\}$, $D = 0$ y $Q \in \{0, 1\}$, manteniendo el periodo $s = 7$ asociado a la periodicidad semanal identificada en el análisis exploratorio. La diferenciación estacional se fijó en cero, dado que el patrón semanal aparecía como una dependencia recurrente estable en los picos de la ACF en retardos múltiplos de siete, y no como una señal que requiriera eliminación mediante diferenciación estacional. Los órdenes P y Q se limitaron a los valores $\{0, 1\}$, ya que la estructura de autocorrelación identificada no aportaba evidencia suficiente para justificar la incorporación de retardos estacionales de orden superior.

En total se evaluaron 72 configuraciones mediante una búsqueda exhaustiva en grilla. Para mantener un coste computacional razonable, la selección inicial se realizó sobre los últimos 180 días del conjunto de validación, aplicando predicción *rolling* a un paso, esquema que reproduce de forma realista el uso del modelo en un contexto de previsión secuencial. La Tabla 4.6 presenta las cinco configuraciones con menor MAE de validación en este tramo reducido. Las mejores especificaciones comparten la misma estructura estacional $(1, 0, 1)_7$, lo que resulta coherente con la periodicidad semanal detectada previamente. En la parte no estacional,

las configuraciones mejor posicionadas incorporan principalmente términos autorregresivos de orden uno o dos y no requieren diferenciación ($d = 0$). Este patrón sugiere que, tras incorporar la componente semanal, la serie puede representarse mediante una estructura lineal relativamente parsimoniosa. La configuración seleccionada fue SARIMA(1, 0, 0)(1, 0, 1)₇, al obtener el menor MAE de validación y mantener una formulación compacta, basada en un término autorregresivo no estacional y una estructura estacional de primer orden.

Tabla 4.6: Top 5 configuraciones obtenidas para el modelo SARIMA mediante búsqueda en grilla.

Order (p, d, q)	Seasonal (P, D, Q, s)	Val MAE (MWh)	Val RMSE (MWh)
(1, 0, 0)	(1, 0, 1, 7)	6,497	8,350
(2, 0, 0)	(1, 0, 1, 7)	7,109	9,128
(1, 0, 1)	(1, 0, 1, 7)	7,326	9,280
(2, 0, 1)	(1, 0, 1, 7)	7,372	9,372
(1, 0, 2)	(1, 0, 1, 7)	7,524	9,617

Una vez seleccionada la especificación SARIMA(1, 0, 0)(1, 0, 1)₇, el modelo se evaluó mediante predicción *rolling* a un paso sobre las particiones completas de entrenamiento, validación y prueba. Esta estrategia permite actualizar la información disponible en cada instante y reproduce de forma más realista un escenario de predicción secuencial. Los resultados se presentan en la Tabla 4.7.

Tabla 4.7: Métricas de evaluación del modelo SARIMA (1, 0, 0)(1, 0, 1)₇.

Conjunto	MAE (MWh)	RMSE (MWh)
Entrenamiento	10,874	38,428
Validación	6,225	8,115
Prueba	6,136	7,944

El modelo alcanza un MAE de 6,136 MWh en el conjunto de prueba, frente a los 9,718 MWh obtenidos por el modelo *naïve*. Esta reducción equivale a una mejora aproximada del 36,9 % en MAE. El RMSE también disminuye de 12,289 MWh en el *naïve* a 7,944 MWh en SARIMA, lo que indica una menor exposición a errores de mayor magnitud durante el periodo de prueba. La mejora observada sugiere que la incorporación explícita de la dependencia autorregresiva de corto plazo y de la componente estacional semanal permite representar parte de la estructura temporal que los modelos de referencia simples no capturan. La diferencia entre el RMSE de entrenamiento y los valores obtenidos en validación y prueba mantiene el patrón observado en los modelos de referencia. El conjunto de entrenamiento presenta un RMSE elevado, de 38,428 MWh, pese a un MAE de 10,874 MWh, lo que refleja la influencia de observaciones extremas localizadas en el tramo histórico. En cambio, los errores en validación y prueba son más próximos entre sí, con MAE de 6,225 y 6,136 MWh, respec-

tivamente. Esta estabilidad entre ambos conjuntos indica que la configuración seleccionada conserva un comportamiento predictivo consistente en los tramos más recientes de la serie.

La Figura 4.16 complementa la evaluación cuantitativa al comparar las predicciones *rolling* del modelo con los valores reales del conjunto de prueba. A diferencia de los modelos de referencia simples, la especificación SARIMA incorpora tanto la dependencia autorregresiva inmediata como la estructura semanal previamente identificada, lo que se traduce en un ajuste más preciso a la serie observada.

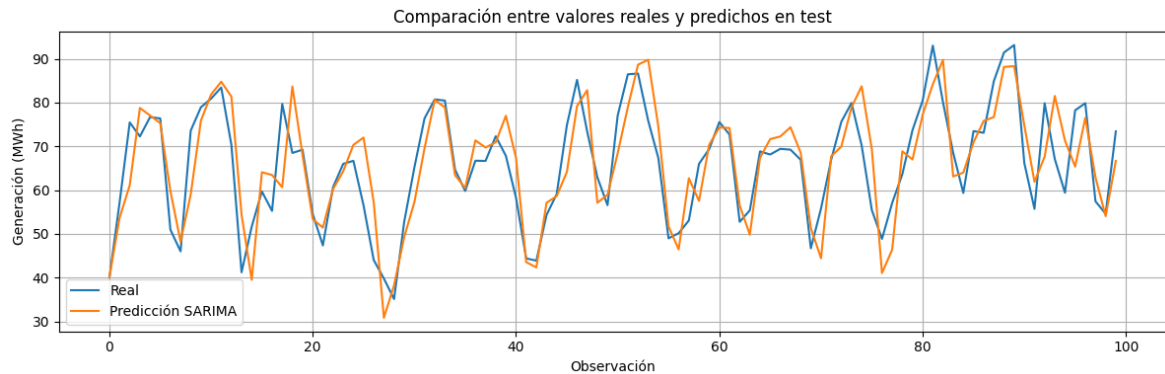


Figura 4.16: Predicciones del modelo SARIMA $(1, 0, 0)(1, 0, 1)_7$ frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.

No obstante, el modelo conserva las limitaciones propias de una especificación lineal. La representación gráfica evidencia que la trayectoria estimada reproduce con fidelidad la dinámica general de la serie en el conjunto de prueba, ajustándose tanto a los cambios ordinarios de la generación como a una proporción considerable de los picos y valles de menor magnitud. Sin embargo, en los días con variaciones más pronunciadas las predicciones tienden a suavizar la respuesta y los residuos aumentan, lo que evidencia una capacidad limitada para representar desviaciones operativas bruscas o cambios de régimen. Esta limitación es consistente con la diferencia observada entre el RMSE de entrenamiento, de 38,428 MWh, y los valores de validación y prueba, próximos a 8 MWh, atribuible a la presencia de episodios extremos en el tramo histórico ya identificados en el análisis exploratorio. Por lo tanto, dicha diferencia no sugiere un problema de generalización, sino una mayor concentración de valores anómalos en el conjunto de entrenamiento. En términos generales, el SARIMA $(1, 0, 0)(1, 0, 1)_7$ representa una referencia estadística competitiva para la comparación posterior. Sus métricas de prueba, con un MAE de 6,136 MWh y un RMSE de 7,944 MWh, muestran una mejora clara respecto al modelo *naïve*, manteniendo al mismo tiempo una formulación interpretable y un coste computacional moderado. Sus principales restricciones se derivan de la linealidad del modelo y de su dependencia de una estructura temporal relativamente estable, aspectos que motivan la evaluación posterior de arquitecturas recurrentes con mayor capacidad para capturar relaciones no lineales.

4.3.3. Modelo RNN

La RNN constituye la primera arquitectura de aprendizaje profundo evaluada en este trabajo. Su incorporación permite analizar hasta qué punto una estructura recurrente básica es capaz de aprovechar las dependencias temporales identificadas en la serie sin recurrir todavía a mecanismos de memoria más sofisticados como los empleados por las arquitecturas LSTM o GRU (Goodfellow et al., 2016). De este modo, la RNN actúa como primer punto de comparación dentro del bloque de modelos neuronales y permite valorar el efecto aislado de la recurrencia sobre la capacidad predictiva. La selección de hiperparámetros se realizó mediante una estrategia de búsqueda aleatoria controlada (*Random Search*) sobre un espacio de configuraciones definido a partir de los resultados obtenidos durante el análisis exploratorio. La longitud de la ventana temporal se exploró en $\{7, 14, 21, 28\}$ días, con el objetivo de evaluar si la red se beneficiaba de incorporar uno o varios ciclos semanales completos dentro de cada secuencia de entrada. El número de unidades recurrentes se fijó en $\{8, 16, 32\}$, abarcando configuraciones de capacidad reducida e intermedia. Asimismo, se evaluaron tamaños de lote (*batch size*) de $\{8, 16, 32\}$ y tasas de *dropout* de $\{0, 0; 0, 1; 0, 2\}$ como mecanismo de regularización destinado a limitar la dependencia excesiva de neuronas individuales durante el entrenamiento (Srivastava, Hinton, Krizhevsky, Sutskever, y Salakhutdinov, 2014).

La combinación de estos valores genera un total de 108 configuraciones posibles. En lugar de realizar una exploración exhaustiva mediante *Grid Search*, se evaluaron 20 combinaciones seleccionadas aleatoriamente con semilla fija con el *Random Search*. Esta decisión se apoya en la evidencia empírica que muestra que, en problemas de optimización de redes neuronales, la búsqueda aleatoria suele alcanzar configuraciones competitivas con un coste computacional significativamente menor, especialmente cuando solo una parte reducida de los hiperparámetros ejerce una influencia dominante sobre el rendimiento final (Bergstra y Bengio, 2012). Todas las configuraciones se entrenaron utilizando el optimizador Adam (Kingma y Ba, 2015), función de pérdida Huber y un mecanismo de *early stopping* con una paciencia de diez épocas sobre la pérdida de validación, restaurando automáticamente los pesos correspondientes al mejor estado alcanzado durante el entrenamiento.

La Tabla 4.8 recoge las cinco configuraciones con menor MAE de validación. La especificación con mejor desempeño combina una ventana de 28 días, 16 unidades recurrentes, tamaño de lote de 16 observaciones y una tasa de *dropout* de 0,1. La predominancia de ventanas de 21 y 28 días entre las configuraciones mejor posicionadas sugiere que la red obtiene beneficio al disponer simultáneamente de varios ciclos semanales completos, resultado coherente con la estructura temporal identificada previamente mediante la ACF, la PACF y los análisis estacionales. Por otra parte, las arquitecturas con 16 unidades recurrentes muestran un comportamiento particularmente consistente, mientras que el incremento hasta 32 unidades no produce mejoras apreciables, lo que apunta a que una capacidad moderada resulta suficiente para representar la dinámica observada en la serie.

Tabla 4.8: Top 5 configuraciones obtenidas para la RNN mediante *Random Search*.

Window	Units	Batch	Dropout	Epochs	Val MAE (MWh)
28	16	16	0,1	50	6,001
21	16	32	0,0	24	6,088
21	8	8	0,0	81	6,159
21	32	16	0,0	19	6,224
28	8	8	0,0	65	6,249

A partir de la configuración seleccionada, se reentrenó el modelo RNN final utilizando una ventana de 28 días, 16 unidades recurrentes, tamaño de lote de 16 observaciones y una tasa de *dropout* de 0,1. La Figura 4.17 muestra la evolución de la función de pérdida Huber durante el proceso de entrenamiento. Se observa una disminución pronunciada en las primeras épocas, seguida de una fase de estabilización progresiva en la que las curvas tienden a valores prácticamente constantes. No se aprecian incrementos sostenidos de la pérdida de validación ni oscilaciones pronunciadas, lo que indica un proceso de optimización estable bajo la configuración seleccionada. La pérdida de validación se mantiene por debajo de la pérdida de entrenamiento durante todo el proceso. Este comportamiento no debe interpretarse como una anomalía del entrenamiento, sino como consecuencia de la estructura temporal de la serie y de la partición cronológica utilizada. Como se mostró en el análisis exploratorio, el conjunto de entrenamiento contiene una mayor variabilidad y episodios atípicos asociados a los registros anómalos identificados en torno a 2020, mientras que el tramo de validación corresponde a un periodo más estable. Por tanto, el modelo se enfrenta a un tramo de entrenamiento más irregular que el utilizado para la validación.

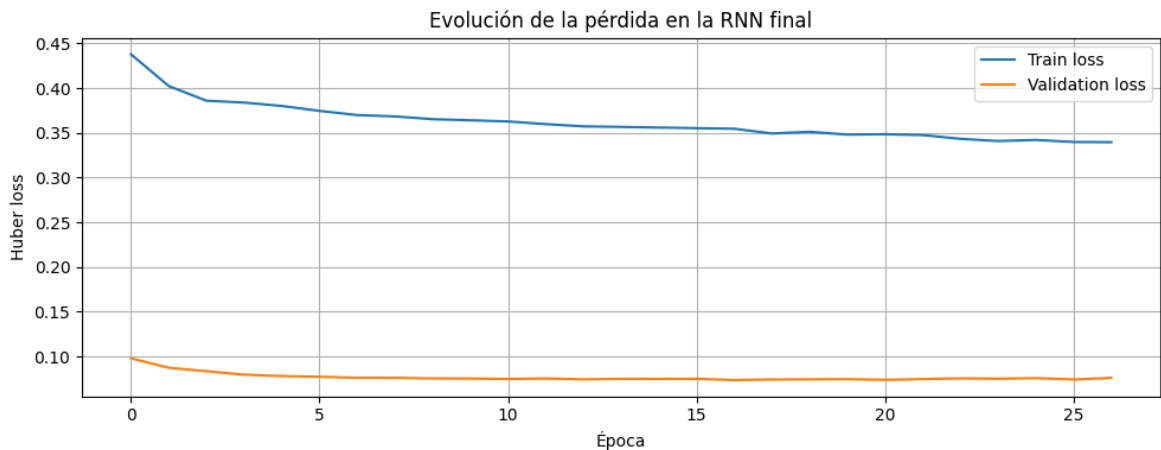


Figura 4.17: Evolución de la pérdida Huber durante el entrenamiento del modelo RNN final. Fuente: Elaboración propia.

La Figura 4.18 presenta la evolución del MAE en entrenamiento y validación, expresado en MWh tras deshacer la escala aplicada mediante *RobustScaler*. El comportamiento es coherente con el observado en la función de pérdida. El error disminuye con mayor intensidad

durante las primeras épocas y posteriormente se estabiliza. En validación, el MAE se mantiene próximo a 6 MWh en la fase final del entrenamiento, sin mostrar una tendencia creciente respecto al entrenamiento. Esta evolución refuerza la interpretación de un ajuste estable, condicionado por una partición temporal en la que el conjunto de entrenamiento presenta mayor variabilidad que el conjunto de validación.

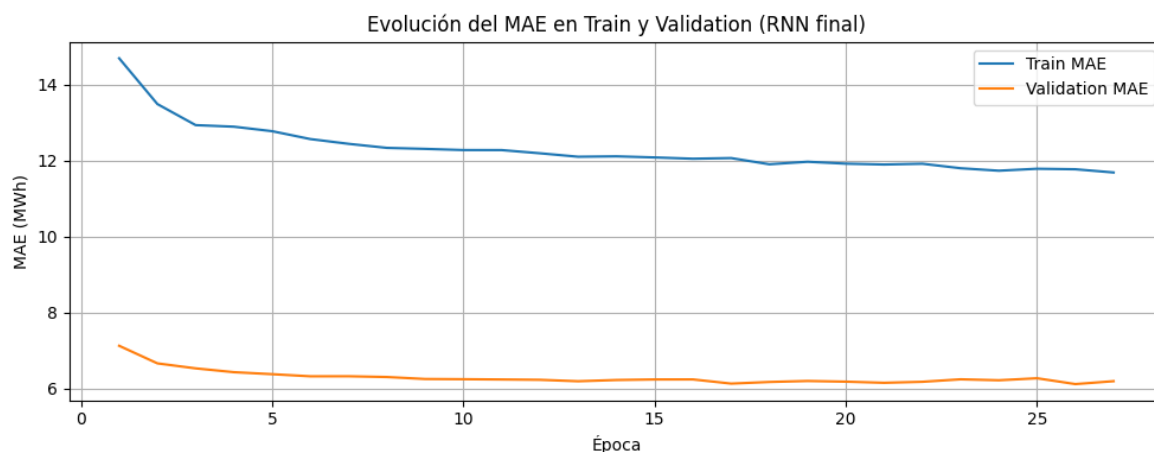


Figura 4.18: Evolución del MAE en entrenamiento y validación durante el entrenamiento del modelo RNN final. Fuente: Elaboración propia.

Las métricas finales del modelo RNN se presentan en la Tabla 4.9. En el conjunto de prueba, el modelo alcanza un MAE de 5,997 MWh y un RMSE de 7,717 MWh. Estos valores mejoran los resultados del modelo *naïve*, cuyo MAE de prueba era de 9,718 MWh, y también superan ligeramente al modelo SARIMA, que obtuvo un MAE de 6,136 MWh. En términos relativos, la RNN reduce el MAE de prueba en aproximadamente un 38,3 % respecto al *naïve* y en torno a un 2,3 % respecto al SARIMA.

Tabla 4.9: Métricas finales del modelo RNN.

Conjunto	MAE (MWh)	RMSE (MWh)
Entrenamiento	11,694	64,250
Validación	6,126	7,786
Prueba	5,997	7,717

La proximidad entre el MAE de validación y el de prueba, 6,126 y 5,997 MWh respectivamente, indica que el comportamiento del modelo se mantiene estable en los dos tramos más recientes de la serie. En cambio, el RMSE de entrenamiento alcanza 64,250 MWh, muy por encima del observado en validación y prueba. Esta diferencia responde a la presencia de observaciones extremas en el tramo histórico de entrenamiento, ya documentadas en el análisis exploratorio, y no a un deterioro del proceso de aprendizaje. La comparación entre MAE y RMSE confirma que los errores de mayor magnitud se concentran en dicho periodo. La Figura 4.19 compara las predicciones de la RNN con los valores reales del conjunto de

prueba. El modelo reproduce adecuadamente la dinámica general de la serie y sigue el patrón semanal observado en los datos. También capta buena parte de las variaciones diarias, aunque tiende a suavizar algunos picos y caídas de mayor amplitud. Este comportamiento es coherente con una arquitectura recurrente simple, capaz de aprovechar la dependencia temporal de corto plazo, pero más limitada para representar cambios abruptos o dependencias más complejas.

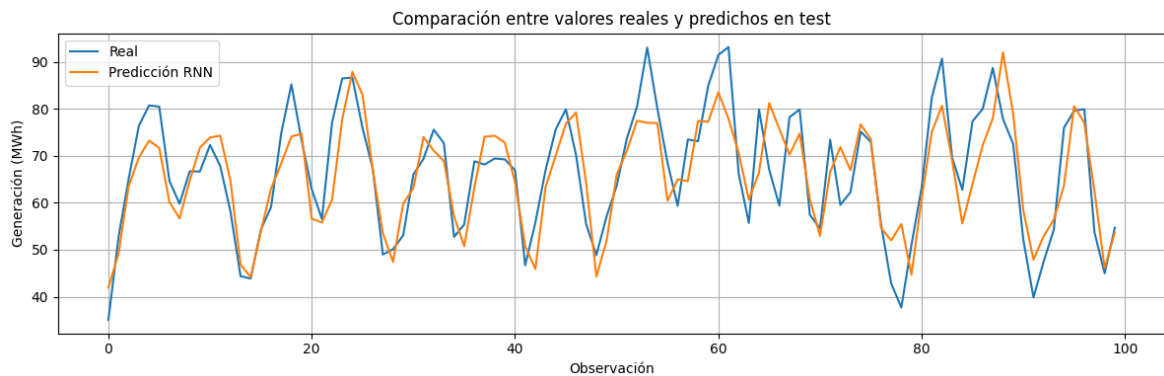


Figura 4.19: Predicciones del modelo RNN frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.

En comparación con los modelos anteriores, la RNN muestra una mejora cuantitativa moderada frente a SARIMA y una mejora amplia frente a los modelos de referencia simples. Este resultado sugiere que la recurrencia básica permite aprovechar parte de la estructura temporal de la serie sin recurrir todavía a mecanismos de memoria más complejos.

4.3.4. Modelo LSTM

La arquitectura LSTM, se evalúa como una extensión de la RNN simple orientada a mejorar el tratamiento de dependencias temporales. A diferencia de la RNN convencional, la LSTM incorpora compuertas de entrada, olvido y salida que regulan el flujo de información dentro de la unidad recurrente, permitiendo conservar o descartar información de forma selectiva a lo largo de la secuencia (Gers, Schmidhuber, y Cummins, 2000; Hochreiter y Schmidhuber, 1997). Su inclusión en este trabajo permite comprobar si una arquitectura con mayor capacidad de memoria temporal mejora el desempeño obtenido por la RNN en la predicción de generación eléctrica a partir de biogás. Para mantener la comparabilidad metodológica, el escalado *RobustScaler*, el espacio de búsqueda, el procedimiento de *Random Search*, la función de pérdida Huber, el optimizador Adam y el criterio de *early stopping* se mantuvieron idénticos a los empleados en la RNN. De este modo, las diferencias observadas en el desempeño pueden atribuirse principalmente al cambio de arquitectura recurrente y no a modificaciones en el protocolo experimental. La única variación estructural fue la sustitución de la capa recurrente simple por una capa LSTM con el número de unidades definido

en la misma grilla de hiperparámetros. La Tabla 4.10 muestra las cinco configuraciones con menor MAE de validación. La mejor configuración utiliza una ventana temporal de 21 días, 16 unidades LSTM, tamaño de lote de 32 observaciones y ausencia de *dropout*. Al igual que en el caso de la RNN, las configuraciones mejor posicionadas emplean ventanas de 21 o 28 días, lo que mantiene la coherencia con la estructura semanal identificada en el análisis exploratorio. Sin embargo, la configuración seleccionada por la LSTM utiliza una ventana algo más corta que la RNN final, lo que sugiere que la arquitectura con compuertas puede aprovechar la información temporal relevante con un contexto ligeramente menor.

Tabla 4.10: Top 5 configuraciones obtenidas para la LSTM mediante *Random Search*.

Window	Units	Batch	Dropout	Epochs	Val MAE (MWh)
21	16	32	0,0	100	6,090
28	8	8	0,0	47	6,115
28	16	16	0,1	78	6,132
21	16	8	0,1	51	6,141
21	16	32	0,2	67	6,280

La ausencia de *dropout* en las dos mejores configuraciones indica que, bajo esta partición temporal y este tamaño de muestra, la regularización externa no mejora necesariamente el desempeño de validación. Además, la configuración ganadora alcanza el límite máximo de 100 épocas sin activación del *early stopping*, lo que apunta a un proceso de aprendizaje más prolongado que en la RNN. Esta diferencia es consistente con la mayor complejidad interna de la LSTM, ya que sus compuertas introducen más parámetros entrenables y pueden requerir más iteraciones para ajustar la representación temporal de la serie. La evolución del entrenamiento muestra un comportamiento estable bajo la configuración seleccionada. La Figura 4.20 presenta la trayectoria de la función de pérdida Huber. Se observa una reducción pronunciada durante las primeras épocas, seguida de un descenso más gradual hasta el final del entrenamiento. La pérdida de validación se mantiene por debajo de la pérdida de entrenamiento durante todo el proceso, patrón similar al observado en la RNN y compatible con la mayor variabilidad del conjunto de entrenamiento frente al tramo de validación. En ningún punto se aprecia un aumento sostenido de la pérdida de validación, por lo que las curvas no muestran indicios de sobreajuste.

La Figura 4.21 muestra la evolución del MAE en entrenamiento y validación, expresado en MWh. El error disminuye de forma rápida en las primeras épocas y posteriormente continúa reduciéndose de manera más lenta hasta estabilizarse en valores cercanos a 6 MWh en validación. La distancia entre ambas curvas reproduce la diferencia ya observada en la pérdida Huber y responde a la distinta complejidad de los tramos temporales utilizados para entrenamiento y validación. Las métricas finales del modelo se presentan en la Tabla 4.11. En el conjunto de prueba, la LSTM alcanza un MAE de 6,128 MWh y un RMSE de 7,973 MWh. Estos valores son competitivos frente a los modelos de referencia, pero no mejoran el

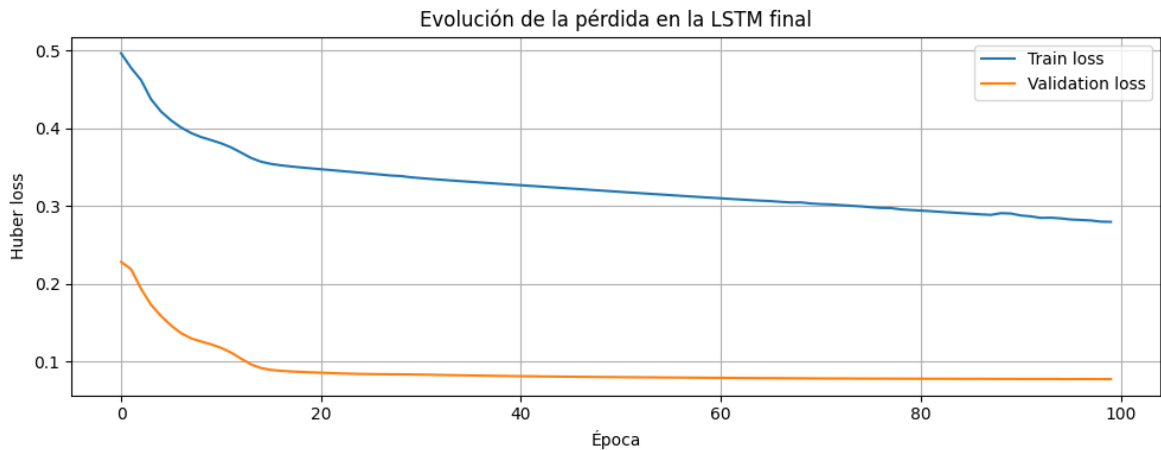


Figura 4.20: Evolución de la pérdida Huber durante el entrenamiento del modelo LSTM final. Fuente: Elaboración propia.

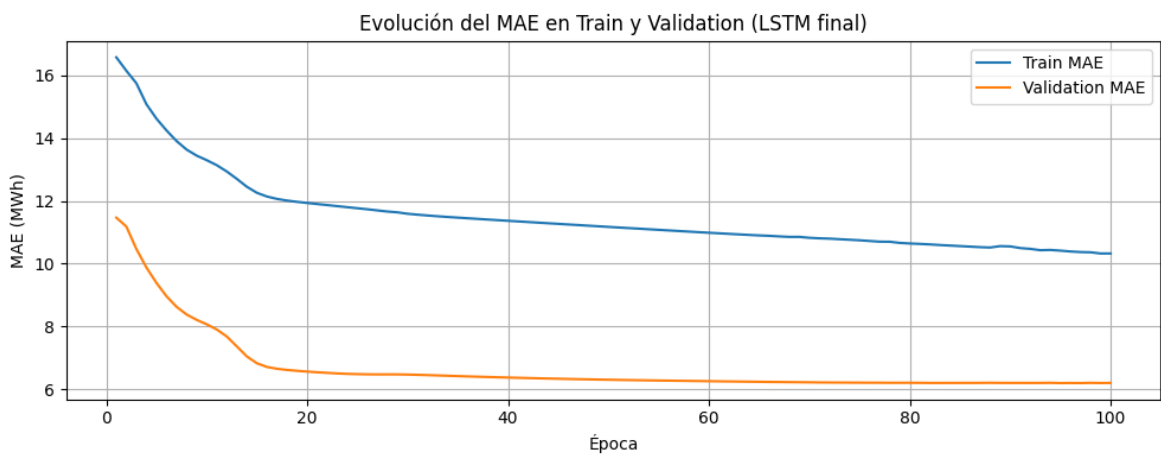


Figura 4.21: Evolución del MAE en entrenamiento y validación durante el entrenamiento del modelo LSTM final. Fuente: Elaboración propia.

desempeño de la RNN, cuyo MAE de prueba fue de 5,997 MWh. La diferencia entre ambas arquitecturas es reducida, lo que sugiere que el mecanismo de memoria de la LSTM no aporta una mejora sustancial en esta serie concreta.

Tabla 4.11: Métricas finales del modelo LSTM.

Conjunto	MAE (MWh)	RMSE (MWh)
Entrenamiento	10,321	49,009
Validación	6,201	7,980
Prueba	6,128	7,973

La proximidad entre los resultados de la LSTM, la RNN y el modelo SARIMA indica que la dinámica predictiva principal de la serie está asociada a dependencias de corto plazo y al patrón semanal ya identificado. En este contexto, una arquitectura más compleja no necesi-

riamente genera una reducción adicional del error. La LSTM muestra un ajuste estable y una buena capacidad predictiva, pero sus ventajas teóricas frente a la RNN simple no se traducen en una mejora clara sobre el conjunto de prueba. La Figura 4.22 compara los valores reales con las predicciones generadas por la LSTM sobre el conjunto de prueba. El ajuste visual resulta similar al observado en la RNN. La arquitectura reproduce adecuadamente la evolución general de la serie y el patrón semanal identificado en el análisis exploratorio, aunque presenta una ligera tendencia a suavizar las variaciones de mayor amplitud. Este comportamiento es consistente con las métricas obtenidas, dado que la diferencia entre el MAE de validación y el de prueba es reducida, 6,201 MWh y 6,128 MWh respectivamente, lo que indica una respuesta homogénea en ambos conjuntos. Aunque la arquitectura LSTM incor-

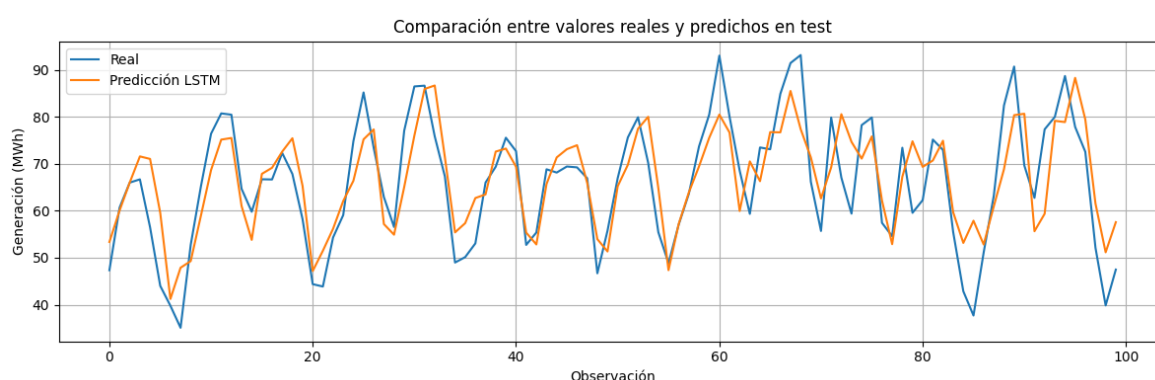


Figura 4.22: Predicciones del modelo LSTM frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.

pora mecanismos de memoria diseñados para capturar dependencias temporales de mayor alcance, esta capacidad adicional no se traduce en una mejora apreciable respecto a la RNN simple. El MAE obtenido en el conjunto de prueba alcanza 6,128 MWh, frente a los 5,997 MWh de la RNN, mientras que el RMSE pasa de 7,717 a 7,973 MWh. Por tanto, dentro de las configuraciones evaluadas, la LSTM mantiene un comportamiento estable y una capacidad predictiva comparable a la de la red recurrente simple, además de una mejora sustancial respecto a los modelos de referencia.

4.3.5. Modelo GRU

La arquitectura GRU, se evalúa como una alternativa a la LSTM con menor complejidad interna. Esta arquitectura, propuesta por Cho et al. (2014), conserva la lógica de compuertas para regular el flujo de información temporal, pero simplifica la estructura utilizada por la LSTM, como se ha explicado en la Sección 3.3. Para mantener la homogeneidad del experimento, el escalado *RobustScaler*, el espacio de búsqueda, el procedimiento de *Random Search*, la función de pérdida Huber, el optimizador Adam y el criterio de *early stopping* fueron los mismos que en los modelos RNN y LSTM. De esta forma, la comparación entre

arquitecturas se apoya en un protocolo común y las diferencias observadas pueden interpretarse principalmente en función del tipo de unidad recurrente utilizada. La Tabla 4.12 presenta las cinco configuraciones con menor MAE de validación:

Tabla 4.12: Top 5 configuraciones obtenidas para la GRU mediante *Random Search*.

Window	Units	Batch	Dropout	Epochs	Val MAE (MWh)
28	8	8	0,0	95	6,032
21	16	32	0,2	67	6,102
28	16	16	0,1	42	6,102
21	16	32	0,0	55	6,115
21	8	8	0,0	86	6,131

La configuración seleccionada combina una ventana temporal de 28 días, 8 unidades GRU, tamaño de lote de 8 observaciones y ausencia de *dropout*. La presencia de ventanas de 21 y 28 días en las mejores configuraciones mantiene la pauta observada en los modelos recurrentes anteriores, donde los contextos temporales de varias semanas resultaron más favorables que las ventanas de una sola semana. Sin embargo, a diferencia de la RNN y la LSTM, la mejor configuración de la GRU utiliza únicamente 8 unidades recurrentes. Este resultado sugiere que la arquitectura puede representar la información temporal relevante con una estructura más compacta, sin necesidad de incrementar el número de parámetros entrenables. La ausencia de *dropout* en la configuración ganadora también indica que, dentro del espacio evaluado, la regularización adicional no mejora el desempeño de validación para esta arquitectura.

A partir de la configuración seleccionada, se reentrenó el modelo final utilizando una ventana de 28 días, 8 unidades GRU, tamaño de lote de 8 observaciones y sin aplicar *dropout*. La evolución del entrenamiento muestra un proceso estable y progresivo. La Figura 4.23 presenta la trayectoria de la función de pérdida Huber a lo largo de las épocas. Tras una reducción rápida durante las primeras iteraciones, la pérdida continúa disminuyendo de forma gradual hasta las etapas finales del entrenamiento, sin oscilaciones pronunciadas ni incrementos persistentes en la pérdida de validación. El proceso se prolongó hasta la época 95, muy próximo al límite máximo establecido de 100 épocas, lo que indica que el modelo siguió encontrando mejoras sucesivas durante gran parte del entrenamiento antes de que se activara el criterio de parada.

La evolución del MAE, representada en la Figura 4.24, resulta consistente con el comportamiento observado en la función de pérdida. Tanto el error de entrenamiento como el de validación disminuyen de forma progresiva hasta estabilizarse en valores cercanos a 6 MWh. La separación entre ambas curvas permanece reducida durante todo el proceso, sin evidencias de divergencia sistemática que sugieran una pérdida de capacidad de generalización. Este comportamiento es especialmente relevante al considerar que la arquitectura seleccionada utiliza únicamente 8 unidades recurrentes. A pesar de esta configuración más compacta,

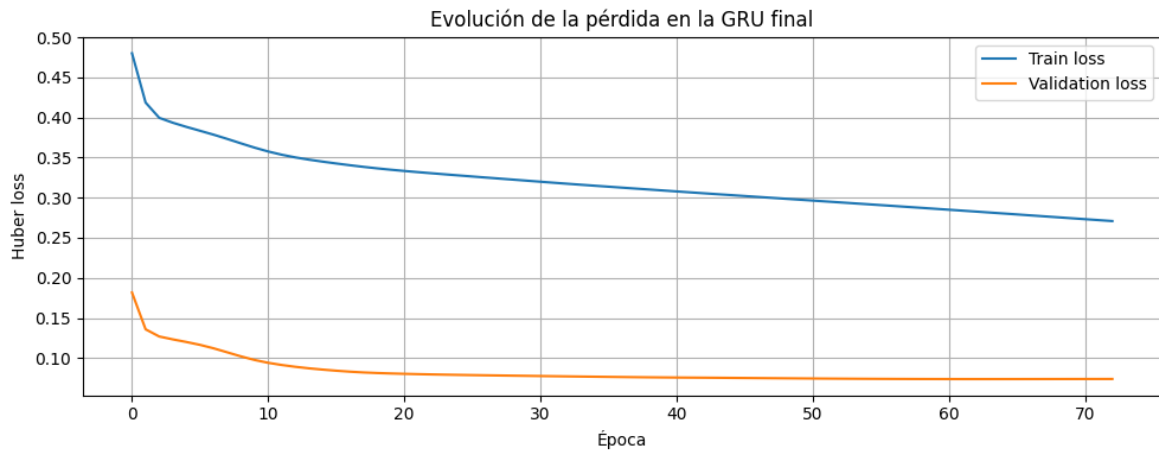


Figura 4.23: Evolución de la pérdida Huber durante el entrenamiento del modelo GRU final. Fuente: Elaboración propia.

el modelo mantiene niveles de error comparables a los observados en la RNN y la LSTM, lo que indica que la información temporal más relevante de la serie puede representarse adecuadamente mediante una estructura recurrente de menor tamaño.

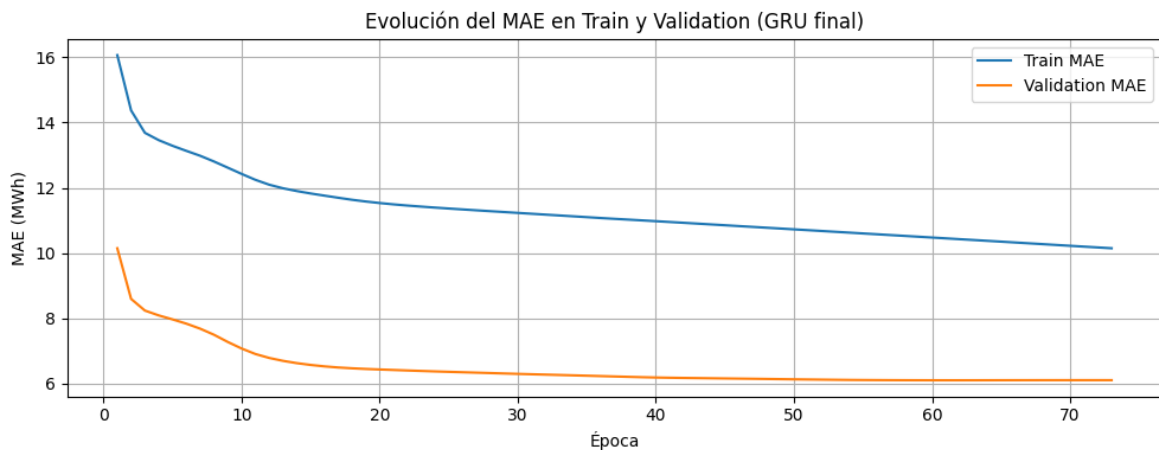


Figura 4.24: Evolución del MAE entrenamiento y validación durante el entrenamiento del modelo GRU final. Fuente: Elaboración propia.

Las métricas finales del modelo GRU se presentan en la Tabla 4.13. En el conjunto de prueba, la arquitectura alcanza un MAE de 5,750 MWh y un RMSE de 7,448 MWh, los valores más bajos obtenidos hasta este punto del análisis. En comparación con la LSTM, la GRU presenta un MAE prácticamente equivalente en entrenamiento, 10,304 frente a 10,321 MWh, pero obtiene errores inferiores en validación y prueba. Esta diferencia indica que la arquitectura más compacta no reduce su capacidad de ajuste sobre el tramo histórico y, al mismo tiempo, mejora su desempeño fuera de muestra.

La comparación de los RMSE de entrenamiento también resulta informativa. La GRU obtiene un RMSE de 49,206 MWh, muy próximo al de la LSTM, 49,009 MWh, y clara-

Tabla 4.13: Métricas finales del modelo GRU.

Conjunto	MAE (MWh)	RMSE (MWh)
Entrenamiento	10,304	49,206
Validación	6,105	7,802
Prueba	5,750	7,448

mente inferior al de la RNN, 64,250 MWh. Esta diferencia sugiere que las arquitecturas con compuertas gestionan mejor los episodios extremos presentes en el tramo histórico, aunque esta propiedad no implica por sí sola una mejora automática en validación o prueba. En el caso de la GRU, sin embargo, la reducción del error se mantiene también en los conjuntos más recientes, con un MAE de 6,105 MWh en validación y 5,750 MWh en prueba. La Figura 4.25 compara los valores reales con las predicciones generadas por la GRU en el conjunto de prueba. El modelo reproduce de forma precisa la trayectoria general de la serie y sigue adecuadamente las fluctuaciones de corto plazo. En comparación con la RNN y la LSTM, la curva estimada muestra una menor tendencia a suavizar los picos y valles, lo que resulta coherente con la reducción observada tanto en MAE como en RMSE.

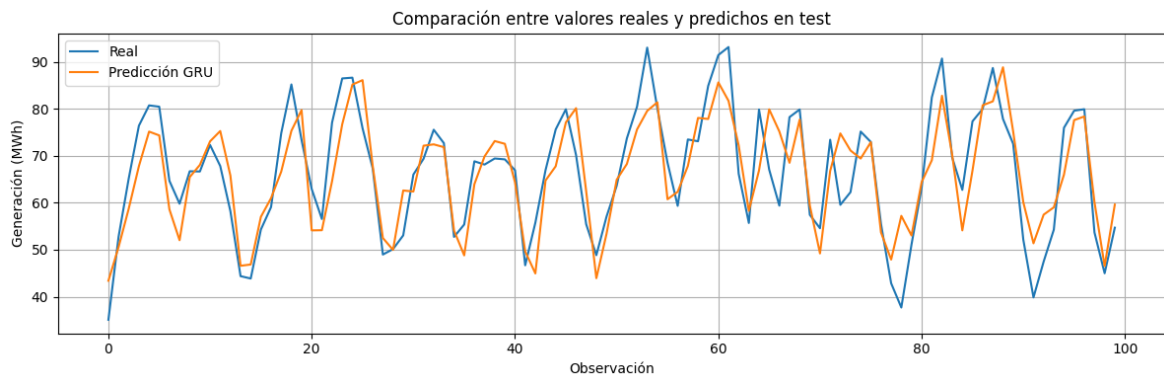


Figura 4.25: Predicciones del modelo GRU frente a los valores reales en el conjunto de prueba. Fuente: Elaboración propia.

El comportamiento del modelo en prueba confirma que la configuración seleccionada ofrece un equilibrio favorable entre precisión y complejidad. Con solo 8 unidades recurrentes, la GRU supera a la RNN y a la LSTM, además de mejorar los resultados del modelo SARIMA y de los modelos de referencia simples. Esta combinación de menor tamaño arquitectónico y menor error fuera de muestra sugiere que, para esta serie, una estructura recurrente compacta con mecanismos de compuerta resulta suficiente para capturar las dependencias temporales relevantes sin incrementar innecesariamente el número de parámetros.

4.3.6. Comparativa global de modelos

Una vez evaluados los modelos de forma individual, se integran sus resultados para comparar el desempeño predictivo alcanzado por cada enfoque bajo las mismas condiciones experimentales. Esta comparación permite valorar si el incremento de complejidad metodológica, desde reglas simples hasta arquitecturas recurrentes, se traduce en una reducción efectiva del error fuera de muestra. La Tabla 4.14 recoge las métricas MAE y RMSE obtenidas en los conjuntos de entrenamiento, validación y prueba para los dos modelos de referencia, el modelo SARIMA y las tres arquitecturas de redes neuronales evaluadas.

Tabla 4.14: Comparativa global de modelos según métricas MAE y RMSE (MWh). Fuente: Elaboración propia.

Modelo	MAE <i>Train</i>	MAE Val.	MAE <i>Test</i>	RMSE <i>Train</i>	RMSE Val.	RMSE <i>Test</i>
Naïve	12,112	10,206	9,718	34,127	12,763	12,289
Media Móvil	15,451	11,371	11,021	54,084	13,753	13,387
SARIMA	10,874	6,225	6,136	38,428	8,115	7,944
RNN	11,694	6,126	5,997	64,250	7,786	7,717
LSTM	10,321	6,201	6,128	49,009	7,980	7,973
GRU	10,304	6,105	5,750	49,206	7,802	7,448

El ranking según el MAE de prueba sitúa en primer lugar a la arquitectura GRU, con 5,750 MWh, seguida por la RNN, con 5,997 MWh, la LSTM, con 6,128 MWh, el modelo SARIMA, con 6,136 MWh, el modelo *naïve*, con 9,718 MWh, y la media móvil, con 11,021 MWh. El orden obtenido mediante el RMSE de prueba es muy similar. La GRU vuelve a presentar el menor error, con 7,448 MWh, seguida por la RNN, SARIMA, LSTM, *naïve* y media móvil. La única diferencia relevante entre ambas métricas se produce entre SARIMA y LSTM, ya que SARIMA obtiene un RMSE ligeramente menor pese a presentar un MAE prácticamente equivalente. Esta variación sugiere que ambos modelos tienen un desempeño muy próximo en prueba, aunque con una distribución de errores ligeramente distinta. La comparación entre los dos modelos de referencia muestra que la media móvil queda por detrás del modelo *naïve* en todos los conjuntos. En prueba, su MAE es un 13,4 % superior al del *naïve*, mientras que el RMSE aumenta un 8,9 %. Este resultado indica que, en la serie analizada, la persistencia diaria contiene más información predictiva que el promedio de la semana anterior. La media móvil reduce la variabilidad de la señal, pero introduce un retraso en la respuesta ante cambios diarios de generación, lo que penaliza su desempeño. Por esta razón, el modelo *naïve* se mantiene como la referencia simple más exigente para cuantificar el valor añadido de los modelos posteriores.

La diferencia entre los cuatro modelos más competitivos es reducida. El intervalo de MAE de prueba entre la GRU, con 5,750 MWh, y SARIMA, con 6,136 MWh, es de 0,386 MWh. Esta proximidad indica que las arquitecturas recurrentes mejoran al modelo estadístico estacional, aunque la ganancia adicional frente a un SARIMA bien especificado es moderada. Al mismo tiempo, todos los modelos paramétricos reducen de forma clara el error respecto

al *naïve*, lo que confirma que modelar explícita o implícitamente la estructura temporal de la serie aporta valor predictivo frente a una regla de persistencia simple. En los modelos recurrentes, la distancia entre los errores de validación y prueba permanece acotada. La RNN pasa de un MAE de validación de 6,126 MWh a 5,997 MWh en prueba; la LSTM, de 6,201 a 6,128 MWh; y la GRU, de 6,105 a 5,750 MWh. Esta estabilidad sugiere que el proceso de selección de hiperparámetros no produjo un ajuste excesivamente dependiente del conjunto de validación. En particular, la arquitectura GRU es el único modelo con una mejora apreciable en prueba respecto a validación, lo que refuerza su comportamiento favorable en el tramo final de la serie.

Un aspecto adicional para interpretar la comparación es el comportamiento en el conjunto de entrenamiento. La Tabla 4.14 muestra que GRU y LSTM presentan valores de MAE prácticamente idénticos en entrenamiento, con 10,304 y 10,321 MWh respectivamente. Ambos con mejor ajuste que RNN, SARIMA, *naïve* y media móvil. Esta proximidad indica que las dos arquitecturas con compuertas alcanzan un ajuste similar sobre el tramo histórico. Sin embargo, la GRU obtiene menores errores en validación y prueba, lo que sugiere una mejor transferencia del ajuste aprendido hacia los periodos no utilizados durante el entrenamiento. La comparación frente al modelo *naïve* permite cuantificar el valor añadido de los modelos paramétricos. En términos de MAE de prueba, SARIMA reduce el error en un 36,9 %, la RNN en un 38,3 %, la LSTM en un 36,9 % y la GRU en un 40,8 %. Si la referencia fuese la media móvil, las mejoras aparentes serían mayores, dado que este modelo parte de un desempeño inferior. Esta diferencia muestra que la elección del *baseline* condiciona la magnitud de la mejora reportada y justifica el uso del modelo *naïve* como referencia principal, al ser el umbral simple más exigente en esta serie.

La comparación frente a SARIMA ofrece una lectura más matizada. Las arquitecturas recurrentes mejoran al modelo estadístico, pero las diferencias son reducidas. La GRU, que obtiene el mejor desempeño global, disminuye el MAE de prueba en 0,386 MWh respecto a SARIMA, equivalente a una reducción aproximada del 6,3 %. La LSTM, en cambio, presenta un MAE prácticamente igual al del modelo estadístico, con 6,128 MWh frente a 6,136 MWh. Esto indica que la incorporación explícita del patrón semanal en SARIMA captura una parte importante de la estructura predictiva de la serie, mientras que las redes recurrentes aportan una mejora adicional limitada. La comparación entre arquitecturas recurrentes también muestra que una mayor complejidad no garantiza menor error. La RNN supera ligeramente a la LSTM en el conjunto de prueba, y la GRU, con una estructura más compacta que la LSTM, obtiene el mejor resultado global. Esta pauta es coherente con una serie cuya información predictiva parece concentrarse en dependencias de corto plazo y en el ciclo semanal identificado en el análisis exploratorio. En este contexto, los mecanismos de memoria más complejos de la LSTM no se traducen necesariamente en una ventaja empírica frente a arquitecturas más simples o más compactas. Desde una perspectiva operativa, la selección del modelo no debería basarse únicamente en el menor error de prueba. SARIMA ofrece una

precisión muy competitiva, una formulación interpretable y menor complejidad computacional, por lo que puede resultar adecuado cuando la trazabilidad del modelo sea prioritaria. La GRU, por su parte, presenta el mejor desempeño predictivo y lo logra con una arquitectura reducida, de 8 unidades recurrentes, lo que la sitúa como la alternativa más favorable dentro de las redes neuronales evaluadas.

Capítulo 5

Conclusiones

La transición hacia sistemas energéticos más sostenibles ha impulsado el interés por fuentes renovables capaces de contribuir a la reducción de emisiones y a la valorización de residuos. En este contexto, el biogás ocupa una posición relevante al permitir la producción de energía a partir de residuos orgánicos procedentes de actividades agrícolas, ganaderas, industriales y urbanas. A diferencia de otras fuentes renovables cuya variabilidad está estrechamente asociada a las condiciones meteorológicas, la generación a partir de biogás depende en mayor medida de factores operativos y biológicos vinculados al proceso de digestión anaerobia, como la disponibilidad y composición del sustrato, la gestión de los digestores y el funcionamiento de los equipos de conversión energética. Por ello, su integración eficiente en el sistema eléctrico requiere herramientas de planificación que permitan anticipar la evolución futura de la generación. La predicción de la producción energética asociada al biogás adquiere así interés técnico y estratégico, al facilitar la programación de las instalaciones, apoyar la operación de las plantas y favorecer una gestión más eficiente de los recursos renovables disponibles.

En respuesta a esta necesidad, la literatura reciente ha explorado diferentes estrategias de predicción orientadas a representar la dinámica temporal de la producción energética a partir de biogás. Los enfoques estadísticos tradicionales, como ARIMA, Holt–Winters o SARIMA, ofrecen una primera aproximación interpretable para series con patrones regulares, pero pueden resultar limitados ante relaciones no lineales, cambios operativos o dependencias temporales más complejas. Por ello, varios estudios han incorporado modelos basados en aprendizaje profundo, especialmente arquitecturas recurrentes como RNN, LSTM y GRU, así como variantes híbridas con mecanismos de selección de variables, atención o descomposición multiescala. Esta evolución metodológica muestra que la elección del modelo no depende únicamente de su sofisticación técnica, sino de la estructura temporal de los datos, la calidad del registro disponible, el horizonte de predicción y la variabilidad propia del sistema analizado.

Con base en este contexto, el presente trabajo tuvo como objetivo comparar diferentes

estrategias de predicción aplicadas a la generación eléctrica diaria a partir de biogás en Galicia. Para ello, se construyó una serie temporal regional a partir de registros provinciales descargados del sistema ESIOS, correspondientes a A Coruña, Lugo, Orense y Pontevedra, durante el periodo comprendido entre abril de 2016 y marzo de 2026. La información original fue depurada y agregada a escala regional, obteniéndose una serie univariante de 3638 observaciones diarias expresadas en MWh. A partir de esta base, se desarrolló un análisis exploratorio orientado a caracterizar la distribución, la tendencia, la estacionalidad, la presencia de valores atípicos y la dependencia temporal de la serie. Posteriormente, se evaluaron modelos de distinta complejidad, incluyendo referencias simples como el modelo *naïve* y la media móvil, un modelo estadístico SARIMA y tres arquitecturas recurrentes de aprendizaje profundo, RNN, LSTM y GRU. La comparación se realizó mediante una partición cronológica en conjuntos de entrenamiento, validación y prueba, utilizando MAE y RMSE como métricas de evaluación en la escala original de generación.

El análisis exploratorio permitió identificar una serie con una dinámica relativamente estable en torno a un rango habitual de generación, aunque afectada por una anomalía puntual de gran magnitud concentrada alrededor de 2020. La corrección de este episodio para fines descriptivos permitió observar con mayor claridad la estructura ordinaria de la serie, caracterizada por una distribución unimodal con mayor concentración de valores en torno a 65–70 MWh diarios. Desde el punto de vista temporal, no se detectó una estacionalidad mensual marcada, pero sí un patrón semanal definido, con menores niveles de generación los lunes y domingos y valores más elevados entre jueves y sábado. Asimismo, la media móvil, la descomposición estacional y las funciones de autocorrelación confirmaron la presencia de persistencia diaria y dependencia semanal, junto con una evolución interanual no lineal marcada por una caída inicial hasta 2019, una recuperación entre 2020 y 2023 y una reducción posterior en los años finales. Estos resultados muestran que la serie no responde a una variabilidad aleatoria, sino a una estructura temporal suficientemente definida para ser aprovechada por modelos de predicción.

Los resultados del modelado muestran que la incorporación de estructura temporal mejora de forma sustancial el desempeño predictivo respecto a los modelos de referencia simples. El modelo *naïve* alcanzó un MAE de prueba de 9,718 MWh y superó a la media móvil, lo que indica que la persistencia diaria contiene mayor información predictiva que el promedio de la semana anterior. La especificación SARIMA redujo el MAE de prueba a 6,136 MWh, evidenciando que la representación explícita de la dependencia autorregresiva y de la periodicidad semanal captura una proporción relevante de la dinámica de la serie. Las arquitecturas recurrentes ofrecieron mejoras adicionales, aunque de menor magnitud respecto a SARIMA. La RNN obtuvo un MAE de 5,997 MWh, la LSTM alcanzó 6,128 MWh y la GRU presentó el mejor desempeño, con un MAE de 5,750 MWh y un RMSE de 7,448 MWh. La diferencia entre GRU y SARIMA fue de 0,386 MWh en MAE de prueba, equivalente a una reducción aproximada del 6,3 %, lo que muestra que el modelo estadístico clásico mantiene un desem-

peño competitivo frente a arquitecturas neuronales de mayor complejidad. Esta comparación permite matizar la contribución del aprendizaje profundo en el problema analizado. Aunque la GRU obtuvo el menor error fuera de muestra, la ganancia respecto a un SARIMA bien especificado fue limitada. Este resultado sugiere que la capacidad predictiva de la serie depende en gran medida de patrones temporales relativamente estructurados, en particular la persistencia de corto plazo y la periodicidad semanal, más que de relaciones no lineales de elevada complejidad. En consecuencia, el incremento de complejidad arquitectónica no se traduce de forma automática en una mejora proporcional del desempeño. Desde una perspectiva metodológica, estos resultados respaldan la conveniencia de contrastar modelos avanzados con referencias estadísticas sólidas antes de atribuir a las arquitecturas neuronales una superioridad predictiva general.

Desde una perspectiva aplicada, la predicción a corto plazo de la generación eléctrica a partir de biogás en Galicia puede abordarse con un nivel razonable de precisión mediante modelos capaces de representar la regularidad temporal de la serie, incluyendo tanto aproximaciones estadísticas interpretables como arquitecturas recurrentes compactas. No obstante, la utilidad de estas herramientas debe situarse dentro de un marco más amplio de planificación energética y desarrollo del biogás. En un territorio donde existe una distancia entre el potencial asociado a la disponibilidad de residuos orgánicos y la generación efectivamente desarrollada, la predicción no constituye un fin en sí misma, sino un apoyo técnico para la toma de decisiones. Su valor se materializa al integrarse con estrategias de operación de plantas, planificación de recursos, gestión de sustratos y evaluación de la capacidad instalada. Por tanto, el avance en modelos de predicción debe acompañarse de actuaciones orientadas a superar las barreras económicas, regulatorias, logísticas y sociales que condicionan el despliegue del biogás en la región.

Los resultados obtenidos también abren líneas de investigación futura. En primer lugar, sería pertinente ampliar el enfoque univariante mediante la incorporación de variables explicativas adicionales. Entre ellas podrían considerarse parámetros operativos de las instalaciones, características y disponibilidad de los residuos utilizados como sustrato, información sobre capacidad instalada, incidencias de mantenimiento, almacenamiento de gas o condiciones de operación de los motores de cogeneración. Las condiciones meteorológicas podrían incorporarse como variables complementarias, especialmente en la medida en que afecten indirectamente a la disponibilidad, composición o tratamiento de los residuos, aunque su papel no tendría la misma centralidad que en fuentes renovables dependientes directamente del clima. La inclusión de esta información permitiría avanzar hacia modelos multivariantes capaces de representar con mayor precisión los mecanismos productivos que condicionan la generación eléctrica a partir de biogás. En segundo lugar, futuras investigaciones podrían evaluar arquitecturas de aprendizaje profundo más avanzadas, incluyendo modelos híbridos, redes convolucionales aplicadas a secuencias, mecanismos de atención o arquitecturas basadas en *Transformers*. No obstante, los resultados de este trabajo sugieren que, en una

serie univariante dominada por persistencia diaria y periodicidad semanal, el incremento de complejidad arquitectónica produce ganancias limitadas frente a modelos estadísticos bien especificados. Por tanto, el potencial de estas arquitecturas podría manifestarse con mayor claridad en planteamientos multivariantes, donde la interacción entre variables operativas, características del sustrato, condiciones de mercado y dinámica temporal pueda generar relaciones más complejas que las observadas en una única serie de generación. Finalmente, sería de interés ampliar el análisis hacia horizontes de predicción superiores a un día. La evaluación de predicciones a varios pasos permitiría valorar la utilidad de los modelos en escenarios de planificación energética de mayor alcance temporal, como la programación semanal de la operación, la gestión anticipada de recursos o la coordinación entre producción, almacenamiento y despacho eléctrico. Asimismo, futuras investigaciones podrían comparar estrategias directas y recursivas de predicción multihorizonte, con el fin de analizar la estabilidad del error a medida que aumenta la distancia temporal entre la observación disponible y el valor pronosticado.

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que *ChatGPT* u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Candela Sánchez - Toscano Botas, estudiante de Programa de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “Predicción de la Generación de Electricidad a partir del Biogás en Galicia: Análisis de Tendencias y Proyección a Corto Plazo”, declaro que he utilizado la herramienta de Inteligencia Artificial Generativa *ChatGPT* u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. Brainstorming de ideas de investigación: Utilizado para idear y esbozar posibles áreas de investigación.
2. Referencias: Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
3. Metodólogo: Para descubrir métodos aplicables a problemas específicos de investigación.
4. Interpretador de código: Para realizar análisis de datos preliminares.
5. Estudios multidisciplinares: Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
6. Corrector de estilo literario y de lenguaje: Para mejorar la calidad lingüística y estilística del texto.
7. Sintetizador y divulgador de libros complicados: Para resumir y comprender literatura compleja.

8. Generador de problemas de ejemplo: Para ilustrar conceptos y técnicas.
9. Traductor: Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado *ChatGPT* u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: Junio 2026

Firma: Candela Sánchez - Toscano Botas

Referencias

- Abdulwahab, S., Jabreel, M., y Moreno, D. (2017). *Deep learning models for paraphrases identification* (Tesis Doctoral). doi: 10.13140/RG.2.2.15743.46240
- Bengio, Y., Simard, P., y Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166. doi: 10.1109/72.279181
- Bergstra, J., y Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13, 281–305.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., y Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5.^a ed.). Hoboken, NJ: Wiley.
- Chen, Z., Ma, M., Li, T., Wang, H., y Li, C. (2023). Long sequence time-series forecasting with deep learning: A survey. *Information Fusion*, 97. doi: 10.1016/j.inffus.2023.101819
- Cho, K., van Merriënboer, B., Gülçehre, Ç., Bahdanau, D., Bougares, F., Schwenk, H., y Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. En *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1724–1734). Association for Computational Linguistics. doi: 10.3115/v1/D14-1179
- Chung, J., Gulcehre, C., Cho, K., y Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*. doi: 10.48550/arXiv.1412.3555
- Dittmer, N., y cols. (2021). Time series analysis and forecasting of biogas production in anaerobic digestion processes. *Microorganisms*, 9(9), 1–18. doi: 10.3390/microorganisms9091879
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211. doi: 10.1016/0364-0213(90)90002-E
- European Commission. (2022). *REPowerEU: Affordable, secure and sustainable energy for europe. clean energy factsheet*. Descargado 2026-01-28, de https://ec.europa.eu/commission/presscorner/api/files/attachment/872548/FS_CleanEnergy.pdf.pdf
- European Commission. (2024). *Biomethane*. Descargado 2026-01-28, de <https://energy>

- .ec.europa.eu/topics/renewable-energy/bioenergy/biomethane_en
- Gers, F. A., Schmidhuber, J., y Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451–2471. doi: 10.1162/089976600300015015
- Goodfellow, I., Bengio, Y., y Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press. Descargado de <http://www.deeplearningbook.org>
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., y Schmidhuber, J. (2017). LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222–2232. doi: 10.1109/TNNLS.2016.2582924
- Hochreiter, S., y Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Hyndman, R. J., y Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3.^a ed.). Melbourne, Australia: OTexts. Descargado de <https://otexts.com/fpp3/>
- Hyndman, R. J., y Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. doi: 10.1016/j.ijforecast.2006.03.001
- IEA Bioenergy. (2022). *The role of biogas and biomethane in pathways to net zero*. Descargado 2026-01-28, de <https://www.ieabioenergy.com/blog/publications/position-paper-the-role-of-biogas-and-biomethane-in-pathways-to-net-zero/>
- Instituto Enerxético de Galicia. (2025). *Potencia eléctrica y térmica de origen renovable en galicia*. Santiago de Compostela. Descargado 2026-01-28, de <https://inega.xunta.gal/publicacions>
- International Energy Agency. (2025). *Outlook for biogas and biomethane: A global geospatial assessment*. Paris. Descargado 2026-01-28, de <https://www.iea.org/reports/outlook-for-biogas-and-biomethane>
- International Renewable Energy Agency. (2025). *Bioenergy and biofuels*. Descargado 2026-01-28, de <https://www.irena.org/Energy-Transition/Technology/Bioenergy-and-biofuels>
- IPCC. (2022). Cross-sectoral perspectives. En *Climate change 2022: Mitigation of climate change*. Cambridge: Cambridge University Press. Descargado 2026-01-28, de <https://www.ipcc.ch/report/ar6/wg3/>
- Jeong, H.-T., y cols. (2021). Deep learning-based prediction of biogas production in anaerobic digestion systems. *Water Research*.
- Kingma, D. P., y Ba, J. (2015). Adam: A method for stochastic optimization. En *International conference on learning representations (iclr)*.
- Lim, B., y Zohren, S. (2021). Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A*, 379(2194), 20200209. doi: 10.1098/rsta.2020.0209

- Ministerio para la Transición Ecológica y el Reto Demográfico. (2022). *Hoja de ruta del biogás*. Descargado 2026-01-28, de https://www.miteco.gob.es/content/dam/mitesco/es/energia/files-1/es-es/Novedades/Documents/00HR_Biogás_V6.pdf
- Ministerio para la Transición Ecológica y el Reto Demográfico. (2023). *Plan nacional integrado de energía y clima 2021-2030 (versión actualizada 2023)*. Descargado 2026-01-28, de <https://www.miteco.gob.es/es/energia/planes-estrategias/plan-nacional-integrado-energia-clima.html>
- Olah, C. (2015). *Understanding LSTM networks*. Descargado de <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Oliveira, P., Marcondes, F. S., Duarte, M. S., Durães, D., Martins, G., y Novais, P. (2024). Assessment of LSTM and GRU models to predict the electricity production from biogas in a wastewater treatment plant. En *Good practices and new perspectives in information systems and technologies - worldcist 2024* (Vol. 986, pp. 64–73). Springer. doi: 10.1007/978-3-031-60218-4_7
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Ribeiro, E. F., Marques, A. C., Junior, J. M., de Almeida, C. A., y Campos, F. R. (2019). Holt–winters forecasting for brazilian natural gas production. *International Journal for Innovation Education and Research*, 7(6), 119–129. doi: 10.31686/ijer.Vol7.Iss6.1559
- Rieke, C. (2018). Modeling and optimization of a biogas plant for a demand-driven energy supply. *Energy*, 145, 657–668. doi: 10.1016/j.energy.2017.12.144
- Salamattalab, M. M., Hasani Zonoozi, M., y Molavi-Arabshahi, M. (2024). Innovative approach for predicting biogas production from large-scale anaerobic digester using long-short term memory (LSTM) coupled with genetic algorithm (GA). *Waste Management*, 175, 30–41. doi: 10.1016/j.wasman.2023.12.046
- Sapli, J., Harders, M., y Rauch, W. (2023). Machine learning for quantile regression of biogas production rates in anaerobic digesters. *Science of The Total Environment*, 872, 161923. doi: 10.1016/j.scitotenv.2023.161923
- Scopinho, J. B. C., y Buratto, R. C. S. (2024). Wavelet CNN–LSTM time series forecasting of electricity power generation considering biomass thermal systems. *IET Generation, Transmission & Distribution*, 18(1), 1–10.
- SEDIGAS. (2023). *Estudio de la capacidad de producción de biometano en España*. Madrid. Descargado 2026-01-28, de <https://biometano.sedigas.es/wp-content/uploads/2023/03/sedigas-informe-potencial-biometano-2023.pdf>
- Şenol, H., Çolak, E., y Oda, V. (2024). Forecasting of biogas potential using artificial neural networks and time series models for Türkiye to 2035. *Energy*, 302, 131949. doi: 10

.1016/j.energy.2024.131949

- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., y Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15, 1929–1958.
- Wen, Q., Zhou, T., Zhang, C., Chen, W., Wei, J., Huang, B., y Xiong, H. (2023). Transformers in time series: A survey. En *Proceedings of the 32nd international joint conference on artificial intelligence (ijcai-23)* (pp. 6778–6786).
- Yildiz Dogan, G., Aksoy, A., y Öztürk, N. (2024). A hybrid deep learning model to estimate the future electricity demand of sustainable cities. *Sustainability*, 16(15), 6503. doi: 10.3390/su16156503